

Europe's nemesis at Maastricht

Next month's meeting of European governments at Maastricht is billed as being the gateway to a more coherent European community, but it will be only if governments can suppress their ill temper.

EUROPE'S way of staggering towards continental cohesion was evidently designed by a bad-tempered statesman with a penchant for regular and frequent domestic rows, preferably in public. That is the lesson of the six-monthly summit meetings of the 12 governments of the European Communities (EC), which are technically meetings of what the Treaty of Rome calls the Council of Ministers. Next month's meeting, at Maastricht in the Netherlands, promises to be more than usually stormy. The agenda includes proposals to amend and to strengthen the Treaty of Rome, to which the EC owe their existence. All member states will suffer a further loss of sovereignty, but differently. The ill temper at Maastricht, signalled by what governments are saying in advance that they will and will not accept, is forgivable only in the sense of being understandable. Have the member states forgotten why they joined the European enterprise in the first place?

Comparisons with federations elsewhere, and in particular with the United States, are inexact. The United States owes its existence to the determination of colonials to be free from the hegemony of an intolerant and insensitive colonial power, whence the emphasis in the US constitution on personal liberty (and the continuing nuisance caused by the failure to control the personal use of firearms). The founding fathers delegated to an elected administration responsibility for foreign policy, mutual defence and the protection of personal liberty throughout the union — and then provided an elected legislature to keep the administration in check, in particular by regulating its funds. Although those who negotiated the Treaty of Rome in the early 1950s sought high-mindedly to outlaw internecine wars, their immediate objectives and their constitutional arrangements were much less ambitious than those of the United States two centuries ago.

Heavy weather

That is why Europe repeatedly makes such heavy weather of each new step towards integration. The European Commission, the analogue of the US Administration, administers the Common Agricultural Policy and otherwise does what it can to ensure that intra-European trade is free and that competition is conducted on a level playing-field. With time, the Commission has ingeniously widened its brief; two decades ago, for example, it began regulating environmental pollution on the grounds that, without pan-European regulations, governments might

use nationally lax regulations to give local industries a competitive advantage. But the Commission is the prisoner of the Council of Ministers, which must approve all proposals for community-wide legislation. There is also an elected (and peripatetic) parliament, which is mostly a talking-shop — neither a check nor a balance.

The impending drama at Maastricht stems from the recognition in 1986, more than 30 years after the Treaty of Rome, that the promised level playing-field was a long way off. The Council of Ministers set itself a deadline, the end of 1992, to bring about that state of grace. The Commission has since laboured mightily, and will have smoothed out most of the bumps by the end of next year. But those who agreed to the European Single Act plainly did not think the implications through. How, in particular, can intra-European competition be genuinely free if the currencies in which people settle their bills are different and liable to fluctuation, even devaluation?

Monetary union

That is the logic of what is called European Monetary Union (EMU), the least contentious half of the Maastricht agenda. Community governments will have to subordinate their central banks to a new European Central Bank and to follow monetary and fiscal policies so that European inflation rates 'converge', making possible the introduction of a single currency. Then, the central bank will impose penalties on governments persistently running excessive budget deficits (which would be a trick for financing domestic spending at the expense of taxpayers elsewhere), but neither the meaning of 'excessive' nor the penalties have been defined.

The logic of EMU may be inescapable, yet there are governments trying to wriggle off the hook. The British is the most conspicuous, insisting that a national currency is so much a part of national sovereignty that a commitment to replace it by a European currency can be given only by the British Parliament and only nearer the time. The reservation is necessary to placate the Euro-dissidents among the government's supporters, but nearer the time, the logic will be even more inescapable.

Yet British diffidence on EMU is a pointer to a recurring fault in the way in which the EC make important decisions. There are two complaints. First, the urgency of the push for monetary union is an afterthought, a consequence of the Single European Act. Would it not have



been wiser to anticipate the need in 1986? The collective failure then to appreciate that EMU would be unavoidable in a single market only fuels the suspicions of those who believe that the EC are simply a federalist plot.

Single currency

Second, the terrors (and the outstanding uncertainties) of a single currency could well have been exorcised by deliberate experiment. All EC governments belong to the European Monetary System (EMS) and now (since Britain joined) all have agreed to keep the value of their national currencies within a predefined percentage of a weighted mean. But if EMS had resources of its own (provided as deposits by members' central banks), it would be able to lend them to national central banks under stress, and so be a prototype for EMU's pan-European bank. In half a decade, half the members of the EC would probably have cause to be grateful to it, and thus more likely to sign up for EMU. The European Commission, always ready with glossy brochures advertising that "Europe is a good thing!", should spend more of its energy persuading politicians, by practical demonstration, that the slogan carries weight.

The much more contentious other half of the Maastricht agenda could similarly have been defused by forethought. The plan is for some degree of political union, in part reflecting the determination of Germany not to subordinate its own central bank, the Bundesbank, to a European version that may be less admirably independent of politicians and less zealous about the value of money. France is also keen, but for different reasons, notably the wish to see Germany locked into an international framework. But there is little agreement on what the scope of political union should be, and no wonder: the EC began as an empirical customs, not a political, union.

So far, the preparations for Maastricht have produced a mix of proposals for administration or decision-making from the centre whose only common characteristic is that half a dozen or so governments support them, and which otherwise are a kind of random sample of the kinds of issues to which governments ordinarily pay attention. A scheme for seeking out supranational criminals will probably survive, but not a plan to harmonize judicial systems. But the fiercest arguments will be about foreign affairs and defence. Some insist on the right to independent action, others seek mechanisms for reaching consensus.

This mixed bag of uncoordinated plans will satisfy neither the Euro-sceptics nor the outright federalists. The former, whose scepticism is reinforced by evidence of needless EC interference in national affairs, will have even more to complain of. The federalists will also be disappointed, wringing their hands over the questions likely to be left aside. One obvious omission is immigration policy; the Treaty of Rome insists on the "free movement of labour", which in principle means that an immigrant to any member country is free to look for work in any other. Each member state thus has a vital interest in the immigration policies of its partners.

So would not a common policy, or mechanism for making one, be a prudent investment? Of course. Yet the courageous governments meeting at Maastricht are likely to duck the issue. Governments such as the British (and, to a lesser extent, the French) have eyes on the short-term electoral benefits of seeming to be tough on immigrants, even at a time when there are more people on the move in Europe than ever before. The result will be that Europe's immigration policy will be, in principle, that of its most liberal government, which is hardly what the potentially restrictive governments intend.

This is but one illustration of how the project for political union will yield a hotchpotch of miscellaneous proposals cutting very little ice. Is it too late to hope that the Maastricht meeting might agree to start again, but on a different footing? There are already substantial areas of European affairs crying out for new political agreements. One, for example, is EC policy on higher education (which does not exist) and on research. Students are already free to seek an education at any European university that will offer them a place, and on the same terms as local entrants, but the movement of academics is complicated and often impeded when academics are titularly government employees.

An agreement to abolish such restrictions would be a political triumph, but would have immediate value. And should not the EC pay some attention to the plight of young people in member states that make too little provision for their education?

Research policy

Present policy on research, which costs Brussels more than \$1,000 million a year, is also ripe for reform. The European Commission's role stems from its inheritance of Euratom and from the conviction in the early 1970s that EC research should foster intra-European collaboration. The result is a 'framework' programme whose chief elements are programmes of applied research in fields such as information technology. The Commission will usually meet half the cost of projects involving industrial and other research groups from at least two European states. The programmes are not well run, and their value is debatable. Yet in the circumstances the Single European Act will create, intra-European collaboration should give way to competition, while the Commission should concentrate on the support of basic research. Is that not a political agreement worth reaching?

Europe, in short, would make faster progress if its ambitions were more realistic. The alternative is to turn the problem of political union on its head, borrowing a few leaves from the US or some other such constitution. Can it, for example, continue to make sense that a continent in which people are free to work wherever they can find a job should maintain judicial and penal systems as different as those now littering Europe? Dealing with contemporary problems on the basis of principles that might win general acceptance would make administrative intervention less obtrusive. □

DOE chops funding for high-energy physics

- Panel mulls distributing a 10 per cent cut
- Fate of Main Injector at Fermilab uncertain

Washington

It seems that no area of physical research will be immune from the forced cuts in the research budget of the Department of Energy (DOE). Last week William Happer, director of DOE's office of energy research, met with the High-Energy Physics Advisory Panel (HEPAP) and asked them to recommend how to allocate a 1993 budget that will be 10 per cent less, in real terms, than that of fiscal year 1992. Previously, Happer had met with advisory committees for fusion research and nuclear science and given them basically the same chore (see *Nature* 353, 372; 3 October 1991 & 783; 31 October 1991).

Unlike the fusion and nuclear science communities, each of which was forced to cancel a major project, the high-energy physicists on the panel hope to spread the pain among all their various programmes and avoid shutting anything down completely. "Everybody will suffer at some level," said HEPAP chairman Stanley Wojcicki of Stanford University.

In a letter that should be completed by the end of this week, HEPAP will offer a Happer a list of priorities in high-energy physics for use as a guide in distributing the cuts, Wojcicki said. The top two priorities are to keep university programmes healthy and to maintain research at Fermilab, the major US particle accelerator. It is essential to protect the university programmes, he said, because they have been squeezed severely over the past decade and more cuts could affect the training of future scientists. Fermilab deserves top priority, Wojcicki said, because "it has the highest discovery potential" among

NUCLEAR WEAPONS

Tritium 'peace dividend'

Washington

By scavenging tritium from the nuclear warheads it has recently agreed to destroy, the United States will be able to obtain enough of the element to fill defence and research needs without building a new production reactor for at least two years, the US Department of Energy (DOE) announced last week. This 'peace dividend' of tritium from old weapons is a windfall for DOE, which had been planning to go ahead with a new tritium-producing reactor at either the Savannah River Laboratory or the Hanford Nuclear Reservation. Now plans for the reactor can be considered along with DOE's reorganization of the US nuclear weapons complex, a process that

high-energy physics programmes.

The most controversial question facing HEPAP is what to do about the planned Main Injector at Fermilab. Already, the \$50 million planned to begin construction on the Main Injector has been moved from the 1992 to the 1993 budget, and HEPAP could not agree on whether it should be pushed back even further. The panel is likely to recommend keeping the Main Injector construction funds in the 1993 budget, Wojcicki said, but with the intention of transferring them to other projects if necessary. It would make no sense to start construction in 1993 if funds were going to dry up in the following years, he noted, and six months from now, the high-energy physics budgets after 1993 will be easier to gauge.

The one high-energy physics project that remains sacrosanct is the Superconducting Super Collider (SSC), but that could change. DOE has always promised that building the SSC, which is funded as a separate item from other energy department research, would not hurt funding of other research, and the current cuts in DOE research budgets are giving SSC opponents ammunition in their fight against the \$8,250-million project. Even George Brown, the chairman of the House Science, Space and Technology Committee, is concerned. In a letter sent in late October to energy secretary James Watkins, he wrote, "When the SSC was proposed by the Administration, Congress was assured that it would not be funded at the expense of ongoing science and research programs. This is obviously no longer true."

Robert Pool

is not expected to be finished until 1993.

With a planned reduction of more than 3,000 weapons, requiring an average of up to 4 grams of tritium each, current US disarmament proposals are expected to free up as much as 10 kg of tritium. DOE officials say that should be plenty for the next few years; additional disarmament negotiations are likely to drastically reduce the number of new weapons needed, and research consumes less than 200 grams of tritium per year. Although safety concerns have currently halted all US reactors that can produce tritium, DOE also plans to restart the 'K' reactor at its Savannah River site by the end of the year.

Christopher Anderson

Hunger strike at volcanology institute

A ONE-WEEK hunger strike has won an apparent victory for ten Soviet geologists protesting a proposed move that would have put their fledgling volcanology institute under the direction of a feared and "autocratic" administrator.

The staff members of the newly formed Institute of Volcanic Geology and Geochemistry (IVGG) in Petropavlovsk-Kamchatsky, Russia, began the strike last month to avoid being put under the command of Sergei Fedotov, who has run the Institute of Volcanology for the past 20 years. In an open letter that has been distributed on computer networks to scientists outside the Soviet Union, the striking researchers wrote, "Imposing of co-authorship, hampering of academic career[s], blocking of trips abroad, and . . . firing of dissidents are typical [of Fedotov's] style of leadership."

After years of complaints about Fedotov, the Soviet Academy of Sciences allowed researchers at the Institute of Volcanology the option of moving to a new institute, the IVGG. Forty per cent took the offer, although they have so far been given little funding or access to research facilities.

But last month Georgy Yelyakov, head of the Far East division of the academy, appointed Fedotov to run a commission overseeing the separation of the IVGG from the volcanology institute. As a result of this decision, the strikers wrote, "in less than half a year, Prof. Fedotov has become our immediate supervisor again. The inevitable effect of these decisions will definitely be the end of our last capabilities to do research work."

They ended their protest last week when Yelyakov sent them a cable promising to halt the appointment of Fedotov until a committee could review the situation.

In 1987, an extraordinary investigative article in *Moscow News* described several cases in which Fedotov demoted outspoken critics at the volcanology institute. The article describes the aftermath of a 1984 letter in which 12 institute researchers wrote to the president of the Soviet academy complaining about Fedotov. Three subsequently had their salaries cut, one lost a laboratory, two were denied approval for their research projects and left the institute, and three were dismissed.

Although the strike is over, US researchers who followed the situation say the Soviet scientists are not relaxing. "They've received assurances like this in the past, so they're skeptical," says Thomas Miller, of the US Geologic Survey's Alaska Volcano Observatory. With a letter to US geologists, Miller and fellow researcher C. Dan Miller are hoping to start a fax and telegram campaign in support of the embattled geologists.

Christopher Anderson

Genentech wins in Japan

Tokyo

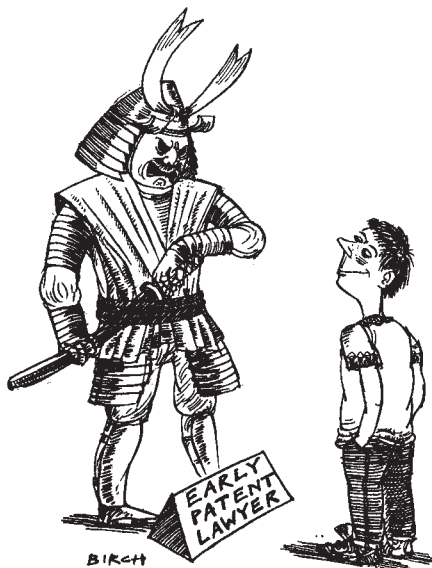
IN a dramatic display of enforcement of Japan's first court ruling on recombinant DNA technology, bailiffs of Osaka District Court last week marched into the Katata plant of Toyobo Ltd to shut down production and seize supplies of a lucrative heart-treatment drug called tissue plasminogen activator (TPA) that is produced using genetic engineering techniques. The court, in a landmark decision, has ruled that Toyobo's sale and production of the drug infringes the patent rights of the US company Genentech Inc.

Genentech and some US observers have hailed the court decision as a sign that Japan is now serious about protecting intellectual property rights. But Japanese experts on the biotechnology industry say the decision, which runs counter to court rulings against Genentech in the United Kingdom and Germany, is flawed and based on an inappropriate decision by the Tokyo patent office.

Although TPA has not turned out to be the blockbuster drug of the biotechnology industry that some predicted, sales of the high-priced drug in Japan have been running at about ¥750 million (\$5.8 million) a month since the drug was put on the market in May. Individual doses (40 mg) of the drug, which helps to dissolve blood clots, are priced at ¥315,000 (\$2,250) each, comparable to the price in the United States.

The market is shared by several companies. Mitsubishi Kasei Corporation, in partnership with Tanabe Seiyaku, and Kyowa Hakko have licensed TPA technology from Genentech, while Daichi Pharmaceutical, in partnership with Toyobo, has been using technology from Genzyme Corporation of the United States.

Long before these companies started production, Genentech filed suit in the Osaka district court in 1987 alleging patent infringement by Toyobo (see *Nature* 333, 587; 1988). And the court has now ruled that the Genzyme technology used by Toyobo does infringe Genentech's Japanese TPA patent, which was awarded by the Tokyo patent office in January.



"We are pleased with the Osaka court's forceful action in protecting our intellectual property rights," says G. Kirk Raab, president and chief executive officer of Genentech. "If this type of fair treatment is displayed in similar cases throughout the appeals process, then Japan will indeed be signalling the worldwide biotechnology industry that meaningful protection, backed by strong remedies, is honoured in

Japan".

But Mitsuru Miyata, editor of the influential newsletter *Nikkei Biotechnology*, says that the TPA court decision is not a good example of recent attempts by Japan to improve its patent system. He points out that Genentech's original TPA patent application had an error in the N-terminal region of the amino acid sequence but the Tokyo patent office allowed the US company to modify its application and remove the part in error while retaining the key part of the amino acid sequence that covers the active part of the TPA molecule. He says the error in the N-terminal region indicates that Genentech's original patent application "seems to be immature" and the Tokyo patent office's subsequent decision to grant the patent followed "bad logic". He also notes that courts in the United Kingdom and Germany rejected TPA patent claims by the US company that were based on the incorrect sequence.

A Toyobo spokesman says there will be an appeal against the court decision. The company is also fighting the Tokyo patent office's award of a TPA patent to Genentech. But a reversal of the decisions is unlikely, observers say, and last week's decision is expected to help Genentech in another court case in Japan against Sumitomo Pharmaceutical Co. Ltd, which has licensed TPA technology from the Wellcome Foundation of the United Kingdom. Sumitomo has yet to win marketing approval from the Ministry of Health and Welfare for its version of the drug.

With the closedown in TPA production, Toyobo and its partner Daichi Pharmaceutical will lose about ¥250 million (nearly \$2 million) a month in sales of TPA; Toyobo has invested more than ¥10,000 million in TPA production facilities and the licence agreement with Genzyme.

David Swinbanks

Solving a dispute, Japanese style

Tokyo

WHILE Genentech of the United States adopted the typical US approach of fighting in the courts to settle its patent dispute over TPA with Toyobo (see above), a group of Japanese companies have solved a patent dispute with Toyobo over another lucrative recombinant DNA drug, erythropoietin (EPO), in typical non-confrontational Japanese style.

Two weeks ago, after inter-company negotiations outside the public eye, Toyobo asked the Ministry of Health and Welfare to withdraw the company's product license approval for EPO. In doing so, Toyobo has shut itself out of a market for EPO in Japan, which is currently worth ¥42,000 million (\$325 million) a year and is rapidly rising. This request to have its license revoked might seem odd to Westerners, but to Japa-

nese it is the only honourable way out of a dispute that Toyobo cannot win.

As in the case of TPA, Toyobo had licensed EPO technology from Genzyme Corporation of the United States. Meanwhile, three other Japanese companies — Kirin beer company, Snow Brand Milk and Chugai Pharmaceutical — have licensing agreements with other US biotech companies and are selling EPO in Japan. Chugai, which has a licensing agreement with Genetics Institute, is in a particularly strong patent position, while Toyobo's position is weakest, observers say. Hence Toyobo's decision to withdraw.

The EPO case also draws attention to the enormous market for this drug in Japan. EPO is a kidney cell glycoprotein that stimulates red blood cell production. It is used to treat patients with chronic renal (kidney) failure.

Because brain death is not recognized in Japan and, as a result, organ transplants are rare, there is a rapidly growing population of patients with chronic renal failure in Japan who are unable to get kidney transplants and who have to get regular (and expensive) dialysis treatment with EPO to survive. The current population of patients is about 100,000 and the number is rising by about 10,000 annually.

The market for EPO in Japan is expected to rise to about ¥70,000 million (\$540 million) per year in the next few years before levelling out because of an expected fall in price of the drug. This situation contrasts with countries such as the United Kingdom where kidney transplants are freely available and the EPO market is negligibly small in comparison.

David Swinbanks

Regulation cost concerns

London

BRITISH biotechnologists are concerned that proposed genetic engineering regulations, published last month by the UK government, will place British companies at a competitive disadvantage within the European Communities (EC).

The regulations are designed to comply with two EC directives regulating the contained use and deliberate release of genetically engineered organisms (see *Nature* **344**, 371; 1990). Among other things, the directives say that no environmental releases of engineered organisms will be allowed without the prior approval of national authorities, and propose that, once a product containing an altered organism is approved in one member state, it can be marketed throughout the EC.

But Louis De Gama, executive director of the BioIndustry Association, is concerned that the regulations will place a greater financial burden on researchers and industry than those being drawn up in other EC states. "We would like to know that there's going to be a consistent approach across the EC," he says.

The British government estimates that implementing the new regulations will cost between £3 million and £10 million over the next five years. Most of this money will come from companies and research institutes, which will be charged for consents to market products containing engineered organisms and to release altered organisms into the environment.

Nigel Poole, from ICI's Jealott's Hill Research Centre in Berkshire, believes that the government's estimate is many times too low. According to the UK Department of the Environment (DoE), most consents for environmental releases will cost between £2,000 and £4,000. But as the precise figure will depend on the administrative cost of each application, De Gama says that some consents may cost tens of thousands of pounds, where applications demand more attention by the government's new Advisory Committee on Releases into the Environment. The cost will be particularly crippling for academic researchers, he says. (The DoE plans to charge the same fees to academics as to industry.)

Officials in the environment directorate-general of the European Commission, which drafted the two EC directives, say that the member states will be left to draw up their own charges for consents to release organisms. Britain is one of four EC states so far to have passed laws to comply with the directive on environmental release. Among the others, Denmark and Germany are proposing similar fees for consents.

Peter Aldhous

Oil fields under control

London

THIS week, the last of the 700 Kuwaiti oil wells sabotaged during the Gulf War will be back under control, a mere seven months after the first well-control teams arrived in Kuwait. That period is surprising short, given that most oil-industry experts had predicted at the time the war ended that some wells would remain ablaze until at least March 1992.

Ralph Brown, who heads the Kuwait Petroleum Company's well-control task force, is one expert who was not surprised. He says his own company had always expected to cap 95 per cent of the sabotaged wells by the end of this year. And, he says, "we were able to mobilize more teams and equipment than people originally thought existed".

Brown adds that scientists and engineers from all over the world — with the notable exception of the United States — deluged his task force with well-meaning, but impractical, technical solutions (mostly based on the misperception that the main problem posed by a blazing oil well is the

fire itself, rather than the uncontrolled flow of oil and gas). The US scientific community was saved from a similar flurry of wasted effort, Brown says, by a symposium held in Washington in April (see Commentary, p. 11), which gave experts a chance to explain to interested US scientists the nature of the problems in Kuwait.

Although the Kuwaiti wells were eventually brought under control using tried-and-tested oil industry technology, Brown says the Washington symposium did produce some "first-rate ideas" that will be extremely useful in the next task faced by Kuwait — clearing the many thousands of unexploded mines and other munitions left from the conflict. In particular, Brown believes that a mine-clearing harrow, developed by William Wattenburg, an independent physicist and radio broadcaster, together with a team from the Lawrence Livermore National Laboratory in California (see p.13), will be a vital tool in the battle to remove the thousands of cluster bomblets now lying in the shallow waters off Kuwait's beaches.

Peter Aldhous

EC RESEARCH

Brite's prospects dim

Brussels

THE European Communities' (EC) centrepiece programme for industrial research, already more than a year behind schedule, faces a further delay. Many of the 12 EC member states are objecting to the European Commission's specific outline for the Brite-EuRam programme, drafted by Filippo Pandolfi, Commission vice-president in charge of research. Pandolfi plans to channel a proportion of the 660-million-ECU (about \$800 million) Brite-EuRam budget into a series of 'targeted projects' aimed at particular industrial sectors.

Brite-EuRam grants, which must be matched by funds from industry, are designed to support a broad range of projects, including research into advanced materials, new manufacturing technologies, raw-materials processing and recycling. Most of the member states, including Britain, argue that EC research money should not be used to benefit particular industries — and so are opposed to Pandolfi's targeted projects, which include specific schemes for the textiles and shipping industries as well as to develop 'clean car' technology. France (where support for particular industrial sectors is an accepted ingredient of government industrial policy) and several of the Mediterranean states are expected to back Pandolfi. If the dispute cannot be resolved quickly, Commission officials say that the entire Brite-EuRam programme will be stalled for three months.

Peter Aldhous

SEMICONDUCTOR RESEARCH

Retrenchment at JESSI

London

FACED with the failure of the European Commission to provide the funds it had promised, the Joint European Submicron Silicon Initiative (JESSI) is cutting back its research programme and focusing on a smaller number of central 'flagship' projects. JESSI spokesman Klaus Knapp, from the German company Siemens, says that JESSI's budget for 1992, at around 400 million ECU (about \$490 million), will be 25 per cent less than the figure expected in 1989, when the initiative was launched.

JESSI is the European equivalent of the US Sematech consortium, bringing together European semiconductor manufacturing companies to work on collaborative research and development projects to help the European industry counter the growing dominance of Japan. One quarter of JESSI's budget was supposed to be paid by the European Commission, a matching sum by individual European governments, and the other 50 per cent by the industrial partners in JESSI projects. But Knapp says that the Commission has provided only one-third of the sum it agreed to pay.

Since 1989, JESSI has supported more than 70 different research projects. Some redesign of the JESSI programme was always planned for 1992, but the current financial difficulties will force more drastic changes than expected. The new flagship projects will include work on lithography, high-definition television, digital audio broadcasting and digital cellular telephones.

Peter Aldhous

Showpiece in transition

Zürich

USUALLY, when one of an institute's researchers wins a Nobel prize, it is a time for some gentle back-slapping — the ultimate public relations boost for any research centre. And at first sight, the Swiss Federal Institute of Technology (ETH, or Eidgenössische Technische Hochschule) in Zürich seems no exception. At the entrance to the research tower housing the laboratory of Richard Ernst, this year's chemistry laureate (see *Nature* **353**, 689; 24 October 1991), a handwritten placard, placed by his ETH colleagues, proclaims Ernst's Nobel success.

But the celebrations at ETH may be relatively short-lived, as its professors struggle to come to terms with the most radical overhaul of ETH's research enterprise in the institute's 136-year history. Over the past two years, the 300 professors and 2,500 scientific and technical staff responsible for research at ETH have been directed into 19 new research departments (biological sciences, chemistry, Earth sciences, forestry, pharmacy and the like). The new system has replaced a staggeringly complex array of more than 130 separate research units (mostly based on laboratories led by a handful of collaborating professors), each accountable directly to ETH's president.

At an institute of ETH's size, spending more than SFr200 million (about \$136 million) each year on research, the previous nineteenth-century organization was pushed beyond its limits, says Heinrich Ursprung, ETH president for 13 years until 1987, and now head of the Swiss agency that oversees federal government spending on basic science. According to

Ursprung, ETH is simply too large for the president to be personally responsible for its entire research programme. In 1987, a team of independent management consultants concluded that some delegation of power was essential if ETH was to retain its position as Switzerland's premier research centre.

But the particular scheme chosen by the board that governs ETH and its sister institute, the Swiss Federal Polytechnic in Lausanne, has met with widespread opposition from ETH's professors. Ernst is dismissive: "It's just nonsense." He believes the 19 departments have placed a layer of bureaucracy between the professors and the president that could undermine the very research excellence on which ETH's reputation is built. Rather than the most productive researchers being rewarded with proportionately more money for basic laboratory supplies and travel (as happened when the president controlled their budgets), Ernst fears that the new department heads will take a more even-handed approach, with most dividing their funds more or less evenly among their professors.

The dissatisfaction of high-fliers such as Ernst with this egalitarian prospect is easy to understand, but, according to Peter Stössel, an assistant to ETH's vice-president for research, Ernst's opposition to the new structure is widely shared. Most ETH professors, it seems, jealously guarded their direct line of communication to the president's office ("a republic of kings" is how one ETH administrator describes the old organization), and now feel devalued by reporting instead to a mere head of department.

This resentment might recede given time, but there are other problems with the new structure that may force further changes. Restructuring ETH into departments seems an obvious development to researchers accustomed to the structure of a typical British or North American research university. But the 19 new research departments at ETH were created alongside an identical number of existing teaching divisions. These divisions are again organized by discipline, but because each is responsible for particular traditional degree course curricula taught to ETH's 11,000 students, their interests do not overlap exactly with those of the new research departments.

To an outsider, the decision to create a parallel structure for research management, rather than unifying ETH's organization for teaching and research, looks like a missed opportunity to reduce bureaucracy. Ex-president Ursprung agrees: in 1987, he favoured the creation of 15 to 20 new departments responsible for both teaching and research. Nevertheless, he accepted the solution adopted by ETH's governing board because, he says, "I felt it was a step in the right direction."

Ursprung believes that too little power has so far been devolved from the ETH president. And Hans Bühlmann, Ursprung's successor as president until his return to mathematics research last year, says that perhaps half of the new research departments have taken a "wait-and-see" approach, and have yet to take over full responsibility for their own research programmes.

More changes are likely — Jakob Nüesch, who joined ETH as president last year, from his post as head of research at the Swiss-based pharmaceuticals company Ciba-Geigy, agrees that some tinkering is

Swiss government plans to boost research

Zürich

THE Swiss federal government, long content to let Swiss-based industry pay for the vast majority of the scientific research carried out within the country's borders, now plans to increase its spending on basic science by 16 per cent a year over the next four years. Last month, the lower house of the Swiss parliament approved a budget that will provide up to SFr2,109 million (about £800 million) for Swiss science from 1992 to 1995.

The new federal government science budget sets the maximum amount of money that will be made available over the next four years, and the parliament is free to spend less than the planning figures when it agrees on each year's level of public expenditure. But Swiss science officials are nonetheless pleased with their political paymasters' new-found discovery of the value of re-

search. Peter Fricker, secretary general of the Swiss National Science Foundation (which is due to receive two-thirds of the four-year budget), describes the federal government's action as "a conscientious effort" to strengthen Swiss basic research.

In return for increasing its spending on science, the federal government is taking a greater role in defining the overall direction of Swiss research. The new planning budget includes more than SFr350 million for six new 'priority programmes' of strategic research, designed to underpin the national economy. Three of these (in biotechnology, informatics and the environment), will be run by the National Science Foundation. The others, in materials science, information technology and optics, are the responsibility of the two main research institutes funded directly by the federal government: the Federal Institute of Technology in Zürich and the Federal

Polytechnic in Lausanne. Fricker expects the science foundation's priority programmes to concentrate on funding university-based researchers, and says that experts from elsewhere will be called in to help with their planning.

The Swiss federal government's decision to boost its spending on science, at a time when science spending in several other European countries is stagnating in the face of economic difficulties, comes in response to a 1990 analysis from the Organization for Economic Cooperation and Development (OECD) showing that, while total Swiss spending on research and development (at around 3 per cent of the country's gross domestic product) was near the average for the developed world, the proportion coming from government was the lowest for all of the OECD countries.

Peter Aldhous

needed — but what the final structure will be is still uncertain. Ernst (who, as president of the research committee responsible for distributing some SFr20 million in federal government grants directly to researchers at ETH, is an influential figure) favours the creation of a handful of schools bringing together like-minded departments, responsible for both teaching and research.

Ursprung, too, favours this idea, provided that the deans of the schools are

given real power to control their own research programmes. "I've always felt one should not try to reinvent the wheel," he says. "If I had the power to decide, I'd adopt the structure of a US private research university." But given the inherent conservatism of its academic staff, ETH's final organization will take some time to evolve — Bühlmann's estimate is four to five years. Any attempt to impose change by "brute force" would fail, he says. "It's the Swiss style."

Peter Aldhous

SCIENCE FAIR

Heureka winds up

Zürich

LAST weekend saw the completion of a unique experiment in science communication, when the Heureka exhibition — an ambitious attempt to explain to the public the current research of Switzerland's scientists — closed its gates after a five-month residency in Zürich. Despite an initially disappointing public response (see *Nature* 351, 597; 20 June 1991), Heureka has achieved its original break-even target of attracting 900,000 visitors. Its director, Georg Müller, can reflect on a commercial job well done, particularly given the opposition to the project from Zürich's city government (largely on environmental grounds), which deprived him of SFr4 million (about \$2.7 million) in expected funding and delayed planning permission until perilously close to the Heureka's opening. But the extent to which the event achieved its aim of communicating the cutting edge of Swiss research to the average citizen remains a matter of debate.

The idea of a research exhibition to celebrate the 700th anniversary of the Swiss federation came originally from Heureka's main sponsors: Switzerland's universities, federal research institutes and scientific academies, and the Swiss National Science Foundation. Under Müller's direction, Heureka has evolved into three exhibitions, rather than one. The scientific organizations' showpiece, a comprehensive display of the best of Swiss research, takes pride of place, housed in a semicircle of eight large tents dominating the Heureka site. But, judging by the distribution of visitors on a cold late-October morning and by sheer wear and tear around the exhibits, a greater draw for most visitors has been the Galileo tower, an exhibition of science through history, and the series of outdoor 'hands on' science exhibits. The latter, including a tightrope-walking bicycle balanced by a heavy weight slung several metres beneath its frame, were introduced by Müller after Heureka

opened, to occupy the younger and more boisterous visitors to the site.

Peter Fricker, secretary general of the Swiss National Science Foundation, seems convinced that the SFr2 million his organization invested in Heureka was money well spent, but accepts that not all of Heureka's visitors have fully appreciated the display of Switzerland's contemporary research. "It is rather high-level," he says. Indeed, the exhibits are surprisingly up to date — the space science section includes data from the Ulysses solar probe, launched one year ago; another exhibit features a solar cell developed by researchers at the Swiss Federal Polytechnic in Lausanne and described in *Nature* only last month (353, 737; 24 October 1991).

Fricker says that one aim of the exhibition was to calm public fear of new technologies. "As far as I can tell, we've succeeded rather well." This has been tackled in a commendably open manner: Heureka's exhibit on genetic engineering, for example, includes a carefully neutral discussion of the possible risks and benefits of the technology, and then leaves visitors to pin their own comments to a notice board. In a country that supports a strong environmentalist movement wary of new genetic technology, Fricker notes with satisfaction the absence of political demonstrations at the Heureka site.

Peter Aldhous

EUROPEAN COLLABORATION

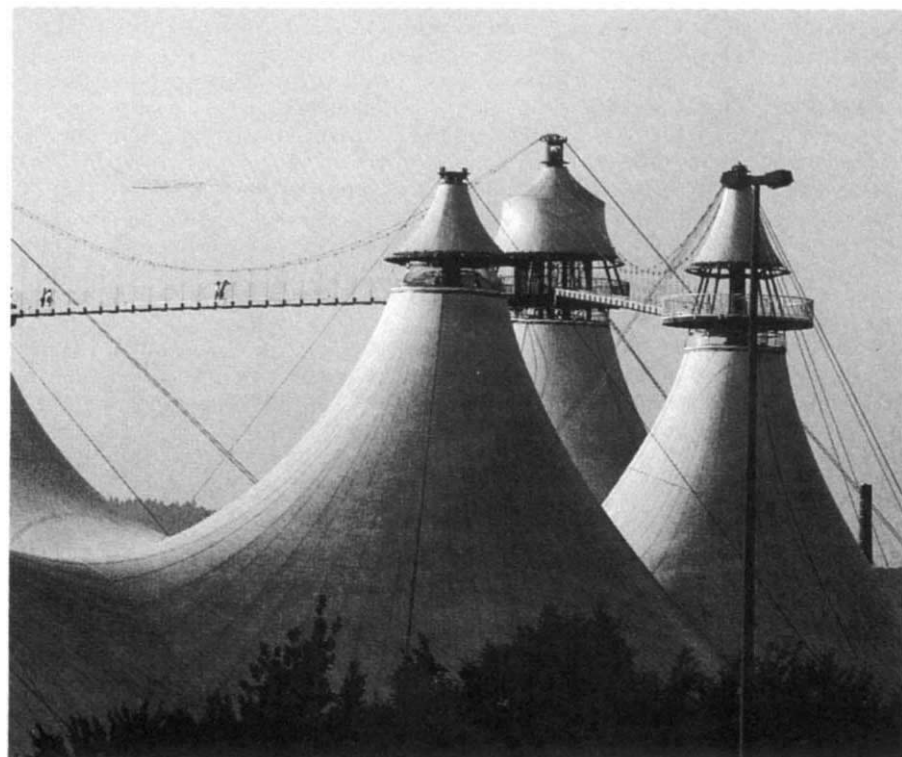
Opening up to the east

London

LATER this month, the beleaguered scientific communities of Czechoslovakia, Hungary and Poland will take a stride towards full integration into the mainstream of European research collaboration. At a meeting in Vienna on 21 and 22 November, ministers from the 19 member countries of COST (European Cooperation in Science and Technology) will welcome the East European states into the organization.

Unfortunately for the new members, COST will provide no badly needed research cash — the organization merely fosters collaborative projects, and has no research budget of its own. But COST membership is an important intermediate step towards eventual acceptance into the collaborative research programmes sponsored by the European Communities (EC). Formed in 1971, COST predates the EC's own research programme. Although COST has now been overtaken by the EC's Framework programme and newer schemes such as the French-inspired Heureka as the principal vehicle for European research collaboration, it still has an important role as a testing ground for collaborative research projects that have yet to win EC backing.

Peter Aldhous



Some of the tents housing Heureka's main scientific display.

Ukraine will close reactors

Moscow

THE report of the official government commission of inquiry into the recent fire at Chernobyl (see *Nature* 353, 690; 24 October 1991) has precipitated a decision by the Ukrainian parliament to close the remaining reactors on the Chernobyl site. The Ukraine is now wondering where its power will come from.

The commission, in a report that echoes the muddle associated with the nuclear accident in 1986, says that the immediate cause of this year's accident was an electrical short-circuit in a multiple-conductor cable linking the generator with the distribution plant some hundreds of metres away. According to the Chernobyl Information and International Contacts Office, damaged insulation was responsible.

The result of the short was inadvertently to cut in the high-voltage switch, making the generator function as an asynchronous motor. The resulting overload on the generator is said to have destroyed the sealing bearings, allowing hydrogen and turbine oil to escape into the machine room. The mixture ignited with a characteristic explosive sound, and two massive torches like gas burners heated up the metal frame of the generator, making it unfit for further use. The roof collapsed, leaving a 2,500-square-metre opening through which rain and snow now fall.

Not until the end of October, more than two weeks after the accident, was the machine room cleared of debris and preparations started for roof repairs. The total damage is estimated at several tens of millions of rubles. Initially, the Soviet Ministry for the Nuclear Industry assessed the accident on its eight-point scale as a Category 1 incident. Under pressure from the official inquiry, it is now ranked Category 2. No accusations have been made against the personnel, still less the fire brigade. But it seems clear that the lesson of the Chernobyl accident in 1986 had not been learned. When the operators began to remove oil from the burning turbo generator, the standby tanks for emergency discharge were found to be filled with waste oil that should have been removed from the site, so that new oil had to be emptied onto the floor.

The commission says that the fire brigade acted bravely in extinguishing the flames in a matter of hours. Just as in April 1986, they had neither special machines, protective gear nor instruments.

The Ukrainian parliament's decision, last week, immediately to shut down the second reactor where the fire happened seems therefore inevitable. But the Ukraine has also decided that the first and third operational units must be shut down "in the shortest possible time and not later than 1993". The notorious fourth nuclear

monster now lies in the sarcophagus but still represents, according to specialists, a dangerous source of radioactive contamination. Ways of finally neutralizing it have not yet been found. Construction of the fifth and sixth power units planned for the site has been halted, and will not be completed.

The overall cost of Chernobyl's assets is more than 1,000 million rubles. Another 500 million rubles went into the construction of Slavutich, the dormitory community of nearly 40,000 power workers 45 km from the nuclear project. The place is within the radioactive zone, and is likely now to become a ghost town. The prospects for the residents are grim. Finding houses and jobs elsewhere will be even more difficult now that the Ukrainian parliament has turned down the Council of Ministers' proposal to put on line power units at other nuclear facilities.

A further difficulty at Chernobyl is that nuclear installations of this size and

sophistication have not previously been closed down and dismantled. A special scientific and technical programme and large sums of money are needed. The Ukrainian Supreme Soviet, fearing that it would not be able to handle the issues on its own, has appealed to the United Nations and to governments of other countries for help.

The parliament must also decide how to make up the power deficit the closure of Chernobyl will create. In 1990, the Ukraine produced 3.7 million tonnes of oil, while importing 58.4 million tonnes from elsewhere in the union. Last month, only eight of the 15 ex-union republics signed the treaty creating an Economic Community, and the Ukraine was not among them. It is believed that this will impede access to oil from outside the Ukraine, but the well-fed Ukraine may have the means to barter with hungry Russia for oil.

Meanwhile, the Ukrainian Parliament has decided to make up for the power losses by emergency construction of a steam-gas facility.

Yuri Kanin

A letter from Dubrovnik

"Today is Wednesday, 23 October.

"For 23 days, the city has been without water, electricity, gas and telephone connections. There are 50,000 citizens and 10,000 refugees of the war within the stone city walls. They are under siege by air, land and sea. All the surroundings have been deliberately burned with phosphorus bombs. Hunger and disease now stalk a city dedicated to culture, the arts and sciences, and to the highest levels of education embodied in its unique Interuniversity Centre.

"The old city of Dubrovnik, catalogued by UNESCO as a World Heritage site, has today witnessed the cruel and disastrous shelling of its ancient centre. The main street, the Stradun or Placa, is devastated and ghostly. The broken glass is in my mind the broken vase of illusion in which we were living. We had been working for a world of responsibility, mutual understanding and support, a world without borders.

"Instead, shells fired by the Yugoslavian Navy have fallen on Lapad, on Gruz, on the hospital at Medarevo, on the Museum Rupe and the School of Music, on Ruder Boskovic street and on the old fortresses of Minceta, Revelin and Bokar. The whole area of the Ploce, from the old harbour to the Hotel Belvedere, has been bombed, as have been the villages of Srebreno, Mlini, Kupari and Plat. The nicest of them all, the town of Cavtat known as the Greek colony Epidaurus, has been seized by federal troops.

"During the week of 20 October, we had planned to hold the fourth scientific congress of the Croatian Biological Society at the Interuniversity Centre. We had planned topics in molecular biology, genetics, evolution, cancer, physiology and health, education in ecology and the timely efforts of the International Council on the Conservation of Nature. But our main topic was to have been the environmental problems of our sea, the Adriatic, to celebrate the centenary of research at the Rovinj Centre for Marine Research.

"We had been honoured by the number of distinguished people from abroad who had accepted our invitation. We were happy to have scientists registered from Slovenia, Italy, Serbia, the United States, Bosnia, Montenegro and Macedonia.

"But now we have not been able to talk about science at all — not about the genetics and conservation of the Mediterranean monk seal, not about the bottlenose dolphins of the Adriatic, not about our plans for saving our sea. Instead, we are sending a desperate SOS on behalf of our potential hosts, who have been let down by Europe and the World. Can nothing be done to save our cultural heritage at Dubrovnik?"

This message is from Professor Drasko Serman of the Biology Department of the Medical School at the University of Zagreb, who is the current President of the Croatian Biological Society.

Making things worse

SIR — So John Maddox thinks (*Nature* 353, 13; 1991) that the AIDS epidemic may have a “silver lining”, does he, because of the work it has allowed Simon LeVay to conduct on the relative sizes of neurons in the hypothalamus of straight and gay men? Apart from this being an insulting way to describe a disease that is probably going to wreck the lives of hundreds of thousands of people, he may also find very few gay men and lesbians who will be quite as pleased at the sort of research he is discussing.

We do not live in a society in which homosexuality is considered as acceptable or normal; rather, gay men and lesbians are harassed, oppressed and generally treated as second-class citizens purely on the basis of our sexuality. In Britain, we are about the only group against which there has been a record of overt direct legislation by the present government and, while more liberal attitudes may be expressed in California, gay men and lesbians everywhere are basically having a very rough time of it. Consequently one would have hoped that scientific establishments would have given some thought to the ethical implications of their research and the social environment it will enter.

The suggestion that Christians would stop defining homosexuality as sinful if it could be shown to be hard-wired is laughably naive. Most Christians traditions maintain that humans are born flawed and to lead a ‘good’ life one must fight to overcome these defects, whatever they are. And frankly, having my status altered from that of being sinful to that of being ill or neurologically abnormal doesn’t seem like much of an improvement to me.

Research of this nature will always be presented in terms of variation from the ‘normal’. I’m sure plenty of readers of the article thought that asking what makes a gay man gay was a perfectly acceptable question but have never asked themselves what makes them straight. The very phrasing of the question reinforces the perception of homosexuality as something alien and inferior, as did the title “Is homosexuality hard-wired?” Research of this comparative kind on social minorities will almost always be used to support the prevailing ideologies and prejudices — one simply has to look at the way IQ data along with neurological and anatomical measurements were used in both the nineteenth and twentieth centuries to support the inferior positions in society given to women and non-Caucasians. And the fact that the current research was conducted by a gay man does not

make it any better — a gay man is just as capable of shooting himself in the foot as anyone else.

The reality of life for a gay man or a lesbian is that our freedom to enjoy the same civil liberties as straight people is very limited as we are constantly confronted by the bigotry of those around us, and work such as that of Simon LeVay will serve only to reinforce this bigotry by presenting society with ridiculously simplistic ‘explanations’ of why we are different from ‘normal’ people. It is also not unlikely that such work will also encourage people to think in terms of a ‘cure’ for being gay, by the use of drugs that act on the neurons in the hypothalamus. This isn’t a fanciful suggestion, as thousands of gay men and lesbians who lived through so-called treatment by such techniques as aversion therapy, electric shock treatment, drug administration and incarceration in mental hospitals can testify. There is more to be considered in scientific research than solely the academic interest that will be generated.

NESSA CAREY

University of Edinburgh,
Edinburgh, UK

Rustum Roy replies

SIR — You take a whole page in News and Views (*Nature* 351, 13; 1991) to advance a case that I am “making a mountain of a molehill” without providing your readers with any factual background on the “issue”. Your attack is full of errors and misinterpretations.

At the request of Tim Beardsley of *Scientific American*, I looked at a paper by Bianconi *et al.* in *Nature*¹ and a piece by Mann² in the same issue. The first described an experiment in precipitating small CdS crystals in an organic gel. Both connected this modest observation by tortured reasoning to “biological composite materials such as bones, teeth and shells”¹ . . . “analogous to those produced by natural biomineralization”². Its significance was alleged to be in . . . “the controlled fabrication of superior composites, ceramics and polymers . . .”² and “in the synthetic mimicking of biominerals”².

As a science policy analyst, I thought these claims to be wildly exaggerated and connected more to a new poorly thought-through push in “biomimetic biomaterials” among US agencies than to any new science. I wrote a two-page policy paper³ and sent it, privately, to agencies, policy makers and journals. The dual thrust of my case was summarized clearly thus: “The case at hand involves two separate errors: (1) exaggeration of the

significance of a particular piece of research by grossly overstating its connection to biomaterials; (2) a total failure to read or refer to the literature by the authors and journal editorial process . . .”.

I continued: “This is NOT a case of cheating, fraud, malice, etc., . . . No one is alleged to have acted out of any but the ‘normal’ processes of science reporting today.” It went on later: “I reiterate: there is absolutely zero suggestion of any intentional error on anyone’s part. But then there’s something very wrong with the system especially because John Maddox and Stephen Mann at *Nature* are so extraordinarily competent.”³

Now compare these facts to your leading article. Your first sentence claims that I charged “plagiarism”, exactly the opposite of what I claimed (that they had not read the literature). And as Editor of *Nature* you go on record that my decrying “exaggeration of the significance of research” and “failure to read the literature” is a “needless fuss about empty issues”. History will prove you as wrong on this as it did when you tried to make a molehill out of the mountain in the Baltimore case.

Third, you achieve an extraordinary feat of hybridizing two scientists, Della M. Roy and myself, by referring to Della M. (“Rustum”) Roy.

RUSTUM ROY

Materials Research Laboratory,
Pennsylvania State University,
University Park,
Pennsylvania 16802-4801, USA

1. P. A. Bianconi *et al.*, *Nature* 349, 315 (1991).
2. S. Mann, *Nature* 349, 285 (1991).
3. Rustum Roy, “Exaggeration of the Potential of Biomimetic Biomaterials,” paper circulated privately, 1–8 (March 1991).

Lice and easy

SIR — Daedalus (*Nature* 352, 761; 1991) proposes an interesting new theory of DNA ‘parasitism’, but misses an important point regarding parasite survival strategy. Lice are highly host- and location-specific, spending their entire life-cycle in a selected niche of their host’s anatomy. On the host’s death then, the louse is not able, as Daedalus suggests, “simply [to] go off and find a new host”, because, with the very rare exception of those like the sheep foot louse¹, transmission of lice from host to host occurs only when the host’s bodies are in contact. In the cuckoo, for example, lice are transferred from one parent to the other at the time of copulation².

MATTHEW PLAYFORD

MASAO KAMIYA

Department of Veterinary Parasitology,
Hokkaido University, Sapporo, Japan

1. Blood, Radostits & Henderson *Veterinary Medicine*, (6th Edn) 955 (Baillière Tindall, London, 1983).
2. Cheng, T.C. *General Parasitology*, (2nd Edn), 639. (Academic, New York, 1986).

Endangered species

SIR — Christopher Anderson's article about the hazards of reintroducing endangered species¹ is timely, given current interest in the role of disease in population dynamics and conservation. But I think he may have spoken with an unusual subset of biologists in coming to the conclusion that most researchers in the area believe that maintaining maximum genetic diversity in wild populations by release of captive animals outweighs the potential costs of simultaneously introducing new pathogens that might extinguish the wild population.

It is very difficult to find evidence that any wild population is in danger of extinction because of lack of genetic heterogeneity². We also know of several species whose populations have increased and are surviving now despite extreme homogeneity. The cheetah *Acinonyx jubatus* is an example. It is so genetically homogeneous that animals from south and west Africa reciprocally accept skin grafts³, yet it survives throughout Africa. In fact, it is beginning to look as if demographic and environmental stochasticity are usually more important problems for small isolated populations than is lack of genetic diversity⁴. And of course, if a wild population is in danger for genetic or demographic or ecological reasons, then in most cases the factors that brought it to that state are probably the real threat, not those biological problems.

To take just one example, the Virunga mountain gorilla population *Gorilla gorilla* numbers about 300 animals, that is, far less than the commonly quoted figure of 500 for large mammals. Even with an effective population size of only

50 breeding animals, more than 100 years will pass before genetic heterogeneity is reduced by 10 per cent. In those 100 years, the local human population will, at current rates of population increase, have doubled, and doubled again. In these circumstances, it would be absurd to count loss of genetic heterogeneity as a threat to the gorilla population⁵; and the same goes for many of the threatened populations of wild species that zoos would try to save by release of captive-bred individuals.

The new findings reported by Anderson¹ confirm what many conservationists have feared all along, namely that the dangers of introduction of disease into wild populations are so high that reintroduction should be a last-ditch solution. Well-run zoos can benefit conservation in many ways; they do not need reintroduction as their grounds for existence, as so many appear to think. Perhaps their main justification is that they encourage the interest in animals that leads to conservation in its rightful (and cheapest) place, namely in the wild.

SANDY HARCOURT

Department of Anthropology,
University of California,
Davis, California 95616, USA

1. Anderson, C. *Nature* **351**, 89 (1991).
2. Allendorf, F.W. & Leary, R.F. in *Conservation Biology. The Science of Scarcity and Diversity* (ed. Soulé, M.E.) (Sinauer Associates, Sunderland, Massachusetts, 1986).
3. O'Brien, S.J., Goldman, D., Merrill, C.R., Bush, M. & Wildt, D.E. *Science* **221**, 459-462 (1983).
4. Lande, R. *Science* **241**, 1455-1460 (1988).
5. Harcourt, A.H. in *Primates. The Road to Self-Sustaining Populations* (ed. Benirschke, K.) (Springer, New York, 1986).

Pseudogenes

SIR — It has been said that humans are a gene's way of making more genes. I propose that the modern streetwise gene's way of assuring its future is the Human Genome Project.

The mapping project will imbue all the sequenced genes with a fitness of 100 per cent. That is, every single sequence will be faithfully preserved and duplicated into eternity. The only selective pressure exerted on a gene will come from scientists who suspect it of causing disease or who seek to use it to cure illness. Such genes might be mutated to accommodate society's healthcare goals. Those genes with interesting biology will out-compete their mundane colleagues for *in vitro* and *in vivo* expression and, more importantly, for scientific inquiry and funding.

With regard to proliferation, genes with widely distributed homology or interesting functions are likely to be copied from central databases onto individual floppies. These sequences might then be expressed on the screens of

countless computers, and even back into animals as transgenes.

As every gene would be individually and directly responsible for its own survival in the database, there would be no need for sex and reassortment to move the gene into a more favourable genetic individual, species or environment. The mixing-and-matching could all be done in computer files.

Interesting rivalries might arise between genes and pseudogenes that had been silenced in the organism. Now, pseudogenes would be as readily expressed on a video screen as the active sequence.

Finally, the very core of natural selection would be imploded. Random mutations could arise by computer error or interestingly enough, cosmic rays, magnetic fields and other forces capable of disturbing storage devices. The ideal genomic database would have a genetic drift of zero.

IBRAHIM A. TANGOREN

David Mahoney Institute
of Neurological Sciences,
University of Pennsylvania
School of Medicine,
107 Johnson Pavilion,
34th and Spruce Streets,
Philadelphia, Pennsylvania 10104, USA

Meanings

SIR — Philological discussions (*Nature* **351**, 179 & **352**, 751; 1991) are interesting in their own right but philology does not dictate or define the meaning of words; meanings are defined by usage. *Mutant* and *recombinant* may have been invented by scientists without a knowledge of language, but these words now have clear meanings. Of greater concern to scientists who care about language should be the overuse of words in ways that destroy meaning. Woe betide anyone who wishes to describe a *parameter*, whose proper meaning is quite clear but which is used instead of a *variable* or a *limit*; and poor *paradigm* has come to mean no more than *example*. The overuse of *demographic* as an adjective applied to the ages and weights of a dozen or so patients studied in a medical research project will soon force epidemiologists to invent another word for true demographic studies. Even worse confusion will come of the increasing habit, especially in US journals, of describing patients entered into studies not as the *sample* but as the *population*. Statisticians had better start looking for new words to distinguish between sample and population statistics.

NEVILLE W. GOODMAN

Department of Anaesthesia,
University of Bristol
and Southmead Hospital,
Bristol BS10 5NB, UK

Cheap big science

SIR — In Sir Mark Richmond's speech "A future for British science" (*Nature* **353**, 379; 1991), he states: "Some research topics are so expensive (or otherwise demanding) that only a few institutions are able to participate in them. Examples are astronomy (at least when it requires large telescopes). . .". In fact, the opposite is true. Large telescopes are now national or international facilities largely equipped with common-user instrumentation, and any astronomer with a good research programme can use them. There are advantages in being in a large department but it is not a necessity.

R. J. TAYLER

Astronomy Centre,
Division of Physics & Astronomy,
University of Sussex,
Falmer, Brighton BN1 9QH, UK

Quenching the wild wells of Kuwait

Richard L. Garwin and Henry W. Kendall

The last of the hundreds of oil-well fires left in Kuwait by Iraq should be capped this week. Although progress has been remarkable, a mechanism is needed to cope with future large-scale disasters.

BEFORE the liberation of Kuwait by US forces and their coalition partners, the government of Iraq detailed about 1,000 military explosive experts and 30–40 engineers, together with the enforced assistance of people who worked in the oil fields, to sabotage and set fire to the 1,000 or so oil wells of Kuwait. Thirty to forty pounds of the military explosive C4 were used with sandbag tamping. Sabotage attempts were made on 732 wells and were successful in damaging and starting oil flows in 690 of them, about 640 of which were also set on fire. Data are scarce, but it seems that something near 2–3 million barrels per day were being lost initially, considerably less than the 6 million barrels per day estimated by Kuwaiti authorities, yet still representing a loss of as much as US\$50 million per day.

The dollar loss was not the only consequence of this act of sabotage: the heavy flows damaged the underground reservoirs; plumes of smoke and combustion products from the flames created serious health and environmental problems, not only in Kuwait but at much larger distances; and lakes of oil developed, more than a metre thick and many square kilometres in extent, creating additional difficulties.

Iraqi land mines and unexploded coalition munitions in Kuwait were expected to add to the hazard and delay, as were administrative impediments. The scale of the disaster was unique: there had never before been more than five wells simultaneously out of control. Earlier experience had shown that an expert well-control company could typically kill and control a burning well in 2 weeks, although months might occasionally be needed. With an estimated 18 teams potentially available, it was believed that 700 wild wells would take several years to control. Numerous ideas were proposed by specialists to deal with the problems of bringing the oil fields under rapid control. But oil wells are far more complex than many realized and the vast majority of the ideas were poorly suited to the challenge. For example, the big problem is not in extinguishing the fires, but in stopping the flow of oil and gas with low risk to the well-

control teams. Moreover, innovative techniques were thought to be required to deal with large numbers of mines in non-combat circumstances. There was initial concern, later shown to be groundless, that wells would be booby-trapped. Mine fields were later found to have been laid outside, but not within, the oil fields.

We, together with Howard Ris and the Union of Concerned Scientists, organised a symposium (later called Desert Quench) in Washington, DC, on 2–3 April this year, to bring together scientists and experts in wells and well control — including several people from the Kuwait Petroleum Company — to address these problems, discuss the usual solutions, and to search for means to expedite control of the wells and deal with challenges posed by the mines.

The problems are interesting, some remaining open despite the speed and

efficiency of the capping-in effort so far. By early October, more than 550 wells had been brought under control, with remaining 140 expected to be capped by the end of this week. Nearly 30 crew were used.

It is evident that we are not the most knowledgeable people about well control, but we write to provide some relevant information, and to indicate how people could contribute to the solution of such problems and similar ones that

The classical approach

The classical and well-practiced control of a wild well involves delivering cubic metres of water per minute to the equipment involved close to the fire, to keep it at operating temperature at a distance at which it would normally melt, and to cool the corrugated-iron shields behind and below which firefighters work to remove debris by the use of caterpillar-tracked bulldozers, cranes and large water-cooled hooks. Getting the fire jetting straight up is almost always the first order of business. For instance, Red Adair in his successful quenching of the Devil's Cigarette Lighter in Algeria (a blazing plume of 10 million m³ of gas per day), first drilled water wells to fill polyethylene-lined reservoirs 10-ft deep and the size of three football fields, about 30,000 m³ of water. (We hope that readers will tolerate a mixture of English and metric units, English being conventional in the oil fields. Conversion is readily achieved as follows: 1 (42-gallon) barrel = 0.159 m³; 1 inch = 2.54 cm; 1 foot = 30.48 cm; 1 pound per square inch (p.s.i.) = 0.0703 kg cm⁻² = 6,980 N m⁻².) Before he snuffed out the fire, Adair cleared the site to table-top neatness and cooled the vicinity of the fire for days to ensure that there would be no re-ignition when the fire was put out. The flame is snuffed out in the classical method by exploding hundreds of kilograms of high explosive carefully shaped inside a thermally insulated oilfield drum supported by a crane at the base of the flame, and detonated from the safety of a trench or burrow at a distance of hundreds of metres.

Adair's biography, *An American Hero*, by P. Singerman (Little,

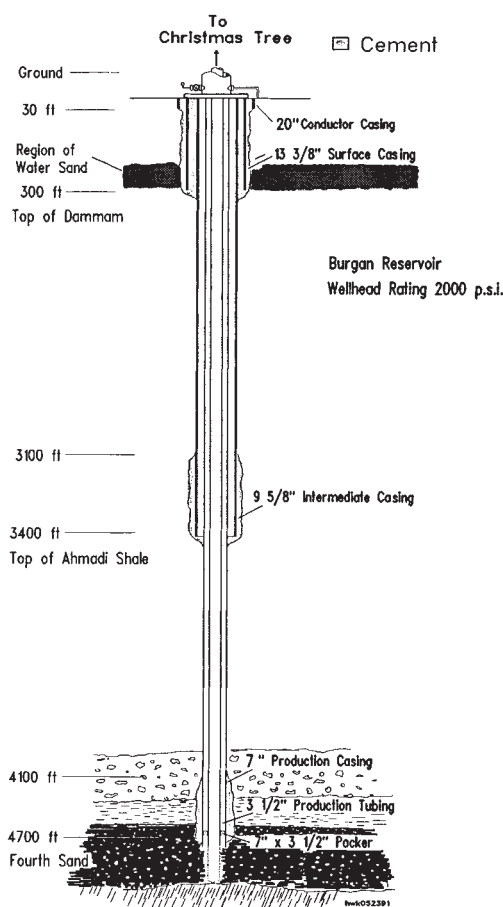


FIG. 1 A typical oil well in the greater Burgan field, south of Kuwait city, not to scale.

Brown, 1990) is a highly readable account of the technology and of the remarkable people who have developed and practiced the demanding profession of control of wild wells, often with an unbeatable safety record. Not one of Adair's team has died in fighting wild wells. But the traditional approach typically takes much time in preparation, and much water; it had been practised at most against five simultaneously burning wells, compared with the hundreds in Kuwait.

The sabotaged wells

The nature of the problems of gaining control of a sabotaged well can be understood by reference to Fig. 1. The well has a production casing of 7-inch diameter extending nearly to the total depth of the well, containing a 3.5-inch production tubing; each tube is of wall thickness adequate for the expected capped pressure. Oil may be produced from the annulus between these two pipes in addition to the production tubing. Figure 1 shows the casing set for a typical well in the Burgan field of Kuwait, indicating the depth of the production zone and other characteristics. The wellhead pressure rating is 2,000 p.s.i., with a test pressure of 4,000 p.s.i.. Some wells have wellhead pressures of more than 6,000 p.s.i. The conductor casing of steel tubing 20–30 inches in diameter is drilled or driven about 30 feet into the ground before well drilling begins to stabilize the drill rig.

The well itself consists of four to seven concentric tubes, including the two described above, that terminate at different depths and may have different pressure ratings. Tubes other than the production tubing are called casings and may be up to 20 inches in diameter. The lower end of each casing is cemented to the formation for a hundred feet or more to prevent flow between otherwise hermetic geological strata. The annulus just inside the largest casing may also be cemented. Annuli may also be closed permanently or temporarily by the use of packers near their lower ends, installed along with the next smaller casing, lining or tubing.

The wells in Kuwait are topped by 'Christmas trees' of plumbing for delivering the oil into the gathering pipeline and for access for well maintenance and control (Fig. 2). The Christmas tree includes many features for control and safety. Blowout protectors, essential to safe drilling, are removed from completed wells, as are the derricks used for drilling and casing. The production tubing and the casings normally hang from flanges in the Christmas tree and, while cemented in at their lower ends, can be under great tension at the surface, the weight of steel — 60 tons or more —

corresponding to about 4 p.s.i. stress per foot of depth. The Christmas tree starts above the floor of a small concrete cellar, typically 2 m in diameter, extending 2–3 m below ground, which reduced the required derrick height during the drilling and casing of the well. Some Kuwait wells are 'sour', yielding substantial H_2S , a gas lethal at concentrations of 200 parts per million, and one that leads

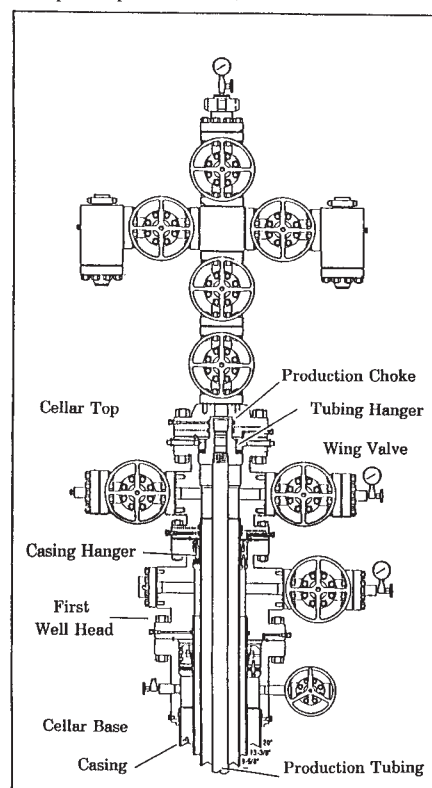


FIG. 2 Wellhead valving and piping, known as the Christmas tree, typical of many in Kuwait. The cellar top is at ground level, the base about 5 ft in the ground. The production tubing is 3½ inches in diameter and successive casings are 9⅝, 13¾ and 20 inches in diameter, respectively.

to hydrogen embrittlement of high-strength steels. The petroleum from various wells in Kuwait has a great range of viscosity, and a gas content reportedly in the range 35–70 m³ of gas at STP per m³ of oil. This high gas content requires a gas-liquid separator under normal production conditions. Sand entrained in the flow at the well foot becomes more abrasive at the higher linear flow speed existing in the two-phase flow near the wellhead; under normal conditions sand-induced erosion of production tubing is monitored and the production tubing replaced when necessary.

After the Iraqi sabotage, it was reported that two-thirds of the wellheads had been blown off and the rest badly damaged. Some of the uncontrolled and burning wells needed only to have their valves closed, but in most cases substantial work had to be done first. In some the valves and wellhead facilities were

removed by explosives above ground level, whereas others had the casing and tubing severed in the well cellar. With pipehangers gone the production casing and the production tubing will drop into the well, damaging it, but not stopping the flow. Many wells are gushing free from unimpeded bores at typical rates of one or two thousand cubic metres per day. Others, whose jets impinge on wreckage or which are venting laterally, are flooding oil at similar rates, which then burns (incompletely) near the ground in great pool fires. The thermal power from a well flowing at this rate is about 300 megawatts. Conical mounds of hardened coke will have accreted around many of the wellheads. The well-known problem of clearing hundreds of tons of drilling rig from the well is absent on occasions like these, of course.

Flow control

With destruction only of the gathering pipeline or of a valve on the Christmas tree above the lowest control valves, the well might be brought under control simply by closing one of these lower valves. But if there are no valves left on a production casing, and no flange, a massive fitting must be assembled to be fitted over the casing after preparation, cooling, and usually after snuffing the flame. This phase of the operation is fraught with danger, for example, from ignition by equipment spark, by lighting from dust or rain storms, or from static electricity.

With eroded pipe, a pit is usually dug to give access to the casing set at a depth at which sound casing can be found. Even if there is no erosion, the casing or casings may need to be exposed and then cut cleanly. The fitting (weighing at least a ton) must be pulled over the geyser in a controlled manner and pulled down over the casing, or clamped around the pipe. Grooved, tapered ramps, called slips, can be inserted between fitting and casing to prevent the large forces that will be present after flow is stopped from pushing the fitting off the casing. For a typical casing of 7-inch diameter and a wellhead pressure of 5,000 p.s.i., this amounts to about 100 tons of force. An appropriate packing or seal is used to retain the fluid after the valve is eventually closed. The high-flow Kuwaiti wells pose particular problems to the traditional approach, as the typical well will spill 1,000 tons of oil a day on the activities at the wellhead between the time the flame is snuffed and the flow is stopped, and may flood the area with poisonous and explosive gas.

Snuffing the flame

Classically, flames are snuffed by the use of high explosive. More recently, as

described at the symposium and used simultaneously by a well-control company in Kuwait, liquid nitrogen has been used to quench the flame in a controllable fashion, with clear advantage over any technique that uses explosives that would imperil people and equipment near the wellhead. Nevertheless, liquid nitrogen has been used only on wells of relatively small flow.

Many have proposed the use of fuel-air explosives instead of high explosives, or the use of Halon gas or even ferrocene to terminate the flame-propagating chemical chain-reactions. Although some of these techniques may be feasible and valuable, it remains true that existing techniques of snuffing flames are cheaper.

Snuffing the fire has traditionally been a prerequisite to stopping the flow, and difficult operations frequently have to precede snuffing. It would be highly advantageous to be able to stop the flow without (or before) snuffing the flame to avoid dealing with the flood of oil and gas that would both hinder work and pose a hazard of explosion. Three approaches are immediately apparent and were discussed by participants of Desert Quench.

(1) Approaching the wellhead with a gathering culvert system, into which the geyser is suddenly diverted without snuffing the flame. This is being developed by several groups.

(2) Tunnelling below the surface to reach the casing set, drilling the production casing (and producing tubing, if necessary) by machines similar to those used at relatively low pressure in domestic distribution systems for hot-tapping an existing gas main, and inserting a stopper(s) adequate to stem or divert the flow. This is not viewed with favour by the drilling industry.

(3) Inserting and fixing a hollow probe or stinger in the bore, which is then valved off. Such a technique has been in use for some time in undamaged wells and was used successfully, early on, in small, damaged wells in Kuwait. It became the technique of choice for the bulk of the well-capping and, with numerous small improvements, has been responsible for the great success of the effort.

In this method, a stinger equipped with appropriate valving is inserted into the flow and restrained by heavy downward force. Powerful pumps are then used to inject drilling mud into the well at a high flow rate until the weight of the mud column halts the flow from the well. Stingers can also be used in wells with flow both through the production tubing as well as through the production casing; the production tubing hanger is cut off and the tubing drops a few feet down, allowing a stinger to be

used in the casing.

Other proposals

Several proposals were introduced and discussed at Desert Quench for speeding up both access to mined and booby-trapped wells and bringing the unconstrained flows of gas and oil under control. A few were unrealistic or unworkable; we discuss some of these briefly here to report the unexpected obstacles that surfaced.

Mine clearing by robust harrows. Conventional mine-clearing techniques are unsatisfactory for the circumstances in Kuwait, particularly because of the higher safety standards traditionally demanded for operations in peace time. The mines are virtually non-magnetic and could be equipped with anti-tamper fusing. Many land antitank mines consist of 10 kg of high explosive packaged in a disk-shaped plastic container, and are capable of driving a hole through the bottom of a heavy armoured tank, and of killing passengers in military or civil vehicles. Such mines can be swept by ploughing, requiring exceptionally well-protected sweep vehicles to prevent injury to their operators. The plough structures rapidly degrade from the detonations. An alternative is to search for the mines on foot by probing with sharp rods. Complex fusing options, which may involve more than one signature, such as pressure or magnetic influence, or which count or use inert periods, make confident clearing uncertain. There have been no satisfactory clearing techniques for mines submerged under a layer of oil.

One potentially valuable approach to mine clearing, proposed by W.H. Watteburg and tested at Yuma proving grounds by Lawrence Livermore Laboratories, uses a sled of chains that incorporates harrow-like blades than can reach to a depth of a foot or so, clearing a 20-ft path. The device digs up mines which either detonate, degrading the sled only slowly, or are entangled in the chains for later destruction. It is drawn by a helicopter at a safe distance. For area-clearing in Kuwait, the sled could be dragged back and forth among several winches, to reduce the cost below the \$3,000 per hour typical of helicopter operations.

Mine clearing with compressed air. Compressed air with sufficient pressure and flow could be used to clear away earth and mines rapidly with a relatively high degree of safety. To illustrate his concept at Desert Quench, S.A. Colgate described a unit with a compressor in the 10,000-hp range; air from a 4-inch nozzle at a pressure in the range 200–400 psi will have a flow of about 5,000 ft³ per min and will develop a reaction force of more than a ton, sufficient to remove

desert soil with great speed down to a depth of from 1–4 ft, as desired. Mounted at the end of a movable boom 30–50 ft long on a tracked vehicle with a protected cab, it might move at 100 ft per min. It would either detonate mines or transport them out of the way along with other debris. Used in conjunction with sled mine sweeping, and with multiple passes over the ground, areas could be cleared to a reasonable level of confidence. It could be used to construct access roads, staging and work areas, and ditches and ponds for either water or oil. The device might dig a pit around a wellhead 60 ft in diameter and 20 ft deep in less than an hour. Not only could such an approach clear mines and booby traps in an area inaccessible because of radiant heat from the fire, it would also remove ground in the wellhead area heated by months of exposure to the flame.

Sand added to the air stream would provide an effective abrasive cut-off tool which would rapidly deal with oil well pipes and valving, reinforced concrete and other materials, facilitating debris removal and nondestructive preparation of the well head. Air-blast excavation and high flow air-abrasive cutting would benefit from increased theoretical and practical study, including the scaling laws applicable to currently available information.

Explosive pipe closure. Ring or cylindrical shell explosives can be used to pinch shut a pipe or multiple annular pipes. The explosive charge or charges must be selected, placed and detonated with more care than, for example, in pipe cut-off. The technique appears unsuitable for oil-well control because of the difficulty of ensuring no further damage. The production tubing and production casing may have suffered wall thinning from stream erosion that is difficult to assess; some of the wells are decades old and the metal of the pipes may have become embrittled from exposure to H₂S. Hence the task of selecting a charge that will confidently close the pipes without severing them becomes difficult, especially as the pipes are suspended under tension. Squeezing pipes shut inside a cemented casing set poses additional difficulty. In some cases, the task may be complicated by the transient pressures expected from the hydraulic ram effect induced by the near-instantaneous halt of the mixed oil-gas fluid in a couple of kilometres of well tubing and, even more important, shock-induced overpressure of the fluid in the pipe or annuli near the explosion that can induce rupture. With the very gassy oil in Kuwait these transient overpressures would be mitigated.

Penetrating bomb. Use of an air-delivered, earth-penetrating bomb to

seal off an injured well 50–100 ft below the surface suffers all the problems and more of explosive pipe closure, owing to the lack of precision with which the bombs can be delivered at depth, and to the inherent one-sided nature of the explosion. Severing a well below the surface converts a difficult situation into a disastrous one, in which the excess pressure in the sub-surface soil exceeds by far the available lithostatic pressure, and leads to the formation of a crater.

Shallow relief well. In cases where the stinger approach is for some reason unworkable, consideration might be given to using a very shallow intercept by a large-diameter, low-pressure, steeply angled well. Connection could be made in conventional manner by explosives at the foot of the relief well, blocking the original well so long as no impediment was posed to the flow of oil out the relief well and into the gathering system. Replaceable lining, resistant to abrasion, would probably be necessary for the relief well. In due time, re-entry could be made through the original well head with several options, including pulling the production tubing from depth and restoring the well.

Tunnelling to the well. Efficient tunnelling techniques could provide a very shallow intercept by a cased, large-diameter tunnel from a point 30–40 m or so from a damaged well to the cemented casing set of the well 30–50 ft below the surface of the ground. The tunnel shaft would be cased and a room mined, shored and cased to provide working space around the casing. To avoid potential explosion hazards from encountering gas or petroleum fumes, or from minor leaks in the apparatus, the tunnel might be kept inerted with CO₂, exhaust gas or nitrogen, with air supplied to masks or hoods for any workers. For the many low-pressure wells, one might bare the outermost casing and attach a saddle to allow drilling under pressure, removal of successive cores, insertion of stoppers and the like — all within the environment of a massive set of pressurized valves and tubes. This technique is essentially foreign to the well-control teams, who emphasize open and quick escape routes, but it might be practised by those experienced in tunnelling, with much help from those experienced in well control.

Flow diverters. It is possible in principle to place a flow diverter over a well (extinguished or not) and conduct the oil–gas mixture into a long culvert-like pipe fitted with one or more sumps that would provide for separation of the gas from the oil, with the gas flared under control. This scheme temporarily halts the production of soot and other pollutants, and recovering the oil. Care would have to be taken to ensure that an

explosion could not develop in the tube and that erosion of the diverter was controlled. The oil could be delivered to plastic-lined holding ponds, dirt tanks, constructed in the desert and using heat-sealed plastic sheets to prevent evaporation and seepage into the sand.

The flow emerging from an uncontrolled well is highly abrasive, in view of the entrained sand and the high velocity to which the gas content carries the fluid on expansion to near-atmospheric pressure. Means are required to retain the integrity of the elbow exposed to this environment; measures such as diamond-coating or massive self-feeding rubber blocks might be useful.

Stove pipes and liquid nitrogen. Cylindrical tubes (3–8 m long) have been used to elevate the flame from a burning well much in the manner of a Bunsen burner. They do not require the cooling screen needed with the burner to prevent flashback; the relative absence of oxygen in the fuel and its high velocity result in the flame base remaining at the top of the tube. A tube hinged along one side, parallel with the tube axis, could prevent its interrupting the fuel stream during insertion but for the lengths so far used this has not been necessary. Liquid nitrogen was used in the first of the damaged Kuwaiti wells, a small one, to have its fire put out. A tube stove pipe was used and the fire extinguished and reignited several times to fine tune the process. Liquid nitrogen can be advantageous where water is in short supply.

Pollution from burning oil. Wells should burn more cleanly if there is better mixing with air (see R. Seitz, *Nature* **350**, 183; 1991), and North Sea drilling operations have long practised the addition of sea water to the oil after the start of free flow of oil to the production platform and into the atmosphere, where the oil is burned to avoid pollution of the sea. The water clearly reduces the atmospheric pollution from the burning oil, and similar approaches could help in Kuwait.

Ducts or driven air can swirl a rising, flaring column of burning fuel. In small flames this produces a narrowing of the column, will raise the level at which burning of the fuel first starts, and appears to promote improved air–fuel mixing and therefore combustion efficiency. If successfully applied to a well fire, such flame control could decrease the radiant heat at the wellhead without the need to extinguish the fire. This would decrease or eliminate the production of pollutants.

Further research into the potential of vortex alteration of large flames would be valuable. On the other hand, if less effort is involved in stopping the flow from a typical well than in providing air or water entrainment, pollution could be

reduced sooner by well control than by modification of the burn.

Gaining control. Shields against the radiant heat of the flame have for decades been made like medieval shields, from corrugated steel, with handles. Spray from water monitors has been used to cool the shields and the fire-fighters, as well as the hot ground. It would be useful to have a light, thin, robust, multi-layer shield that would not be damaged by contact with water or oil. Reflective shields do not long remain reflective in the rain of combustion debris around a burning well. It would also be desirable to have a shield mat that could be placed on the ground to keep shoes from melting or burning.

For wild wells that feed a pool fire instead of a plume fire, control might most readily be achieved in two steps (without first snuffing the flame). First, a surface culvert could be built to approach near the wellhead, with oil diverted into it from the pool by means of a centrifugal pump placed remotely in the well cellar or in an expedient sump. A large lid could then be placed remotely over the wellhead area, excluding air and serving as a crude gas–liquid separator, with the gas taken through a large elevated culvert to be flared (burned) at a reasonable distance.

Conclusions

We have attempted to provide an analysis of the problems and constraints relating to controlling the types of wild wells encountered in Kuwait. We hope that new ways of dealing with such catastrophes will emerge from our discussion. But it is difficult to prepare for a disaster that is likely to happen to somebody else, in a community in which interests are divided and even opposed. In the case of Kuwait, there have certainly been missed opportunities. For instance, if a share in the future production of several wells had been given to one company (and that replicated tenfold), with the proviso that the methods used to control their wells would be available to be practised by others, progress might have been made more rapidly and with higher confidence. As \$50 million a day was being lost in Kuwait, considerable money could have been saved by a fast response to the problem. Not only have important resources been destroyed by the action of Iraq, they have been unnecessarily squandered by an inability to permit technical measures to be implemented to control the wild wells of Kuwait. □

Richard L. Garwin is at the IBM Thomas J. Watson Research Center, PO Box 218, Yorktown Heights, New York 10598, USA; Henry W. Kendall is in the Department of Physics, MIT, Cambridge, Massachusetts 02139, USA.

Metamorphosis of a British laboratory

The British fashion for converting public laboratories into self-governing agencies seems to have helped rescue one of the oldest of them from the torpor into which it was driven in the 1970s.

ONCE upon a time, the National Physical Laboratory (NPL) was Britain's pride and joy among laboratories. Indeed, for a few decades after its creation in 1900, it may have been Britain's only public laboratory of any size and reputation. It was the laboratory at which Watson-Watt invented radar, but even after the Second World War, NPL cut the figure of an intellectual powerhouse, competition from more glamorous places such as Harwell (for atomic energy) and Malvern (for radar and electronics) notwithstanding. In 1957, for example, it startled the British research establishment with a percipient conference on artificial intelligence, complete with an early demonstration of optical character recognition.

Those days have long since gone. The recent history of NPL is a replica of what has happened to British science since the Second World War. In many ways, the laboratory has been lucky; there have been times in the recent past when it may have been saved from extinction only by its responsibility for the maintenance of national standards of length, mass, time and so on — and by fear of the outcry there would have been if such a distinguished place were shut.

Until the end of the 1950s or thereabouts, NPL was buzzing with projects and ideas. Apart from the flirtation with artificial intelligence (which grew into a substantial but largely unproductive programme of computer technology), it then maintained both Britain's largest tank for testing ship models (now spun off into a separate organization) and a substantial programme of aerodynamic research, wind tunnels and all.

Even so, sustained by the then common British belief that an ounce of flair is worth a pound of effort, NPL was stretched too thin over many of the fields it purported to cover. Although large even by present standards, with more than 2,000 people at its peak, the laboratory could pretend at excellence only in some of the areas of its activity.

Then began the period of recurrent reorganization that has played havoc with a host of British organizations. When the old Department of Scientific and Industrial Research was abolished in 1964, NPL was shunted off to the Board of Trade on the strength of its metrology. Now it is a dependant of the Department of Trade and Industry. Never quite knowing where it belonged, who can be surprised that it may

not have known clearly what it was for?

With a little luck, this may now have changed. For just over a year, since July 1990, NPL has been a government agency, which means that it stands on its own feet financially, driven by financial targets set each year by the Department of Trade and Industry, and balancing income (90 per cent of it from the sponsoring department) against expenditure. Small though the outward differences have been so far, this last transformation is probably irreversible. What else could there be except outright privatization (with which the department has been dicker on behalf of another of its laboratories, the National Engineering Laboratory, for the past two years)?

Probably the most noticeable consequence of agency status has been that the 25-hectare site at Teddington, part of south-west London on the left bank of the Thames, looks like a building site. A little to its surprise, one gathers, NPL has persuaded the department to spend £60 million on a new laboratory block. They are already at work threading heating ducts around the two listed buildings that will have to be preserved when the new block goes up.

The most obvious intangible change is the management of NPL. The director, Dr Peter Clapham, his deputy (and controller, scientific services), Mr Andrew Wallard, and the heads of the six divisions form the equivalent of a company board which determines policy and allocates resources accordingly. The influence of the sponsoring department is exercised through a steering board, comprising external people and ministry officials.

What are the benefits? Flexibility is the most obvious. Although the people at the laboratory remain civil servants, NPL can now take on contract research without advance reference to its sponsor, hiring the people to do the work outside the usual constraints on manpower. It can also reward exceptional people with promotions that come a little earlier than might otherwise be the case.

The new ethos is that contracts won independently of government sponsors are valued more highly than the contribution they make to turnover. In the last complete financial year, NPL earned £3.4 million that way — some 7 per cent of its total budget. The obvious advantage is the sense of independence such income engenders. But one member of the staff wryly

noted that this is also a way in which the sponsoring department might eventually reduce its obligations.

A few members of the staff seem more enlivened than threatened by the new opportunities. The micrometrology laboratory, for example, is busily fabricating thickness standards in the micron range for use in the chip industry and has designed a laser interferometer, half the size of a match-box, deliberately so that it can be manufactured at a reasonable cost, in the hope that their partner company will sell a few hundreds rather than a few.

Much in the same spirit, the group using lidar measurements for the measurement of atmospheric pollutants has built itself a mobile laboratory, complete with a twin-dye laser system tunable from the infrared to the ultraviolet, which can be driven to the sites of factories suspected as the sources of, say, hydrocarbon pollution. The group's services are much in demand as environmental regulations bite.

The same group is saddled with the task of pinning down Britain's emission of greenhouse gases, to which end it has (among other things) mounted methane detectors around pens of cattle feeding naturally to measure this source of methane directly.

The standards work also survives, apparently in good shape. One group has built on a single chip an array of Josephson junctions in series that will serve as a standard 1 Volt when popped into liquid helium. The goal of relating all the electrical standards to fundamental constants (electronic charge and mass and Planck's constant) is around the corner, and there is even a scheme for linking them with mass as well.

Yet NPL is no longer distinguishable from a British academic laboratory only by its equipment (traditionally better). Peter Clapham says that he is cheered that the past five years' harvest of recruits have readily taken to the idea that their research is no more valuable than they can persuade their sponsors that it is.

But these are early days. The long-term test will be that of whether the sponsors remain loyal to the standards programme, and whether enthusiasm for winning contracts distracts from thoughtfulness. And meanwhile, the professional unions are worrying that success at NPL may bring in an army of people on short-term contracts, without a traditional career structure.

John Maddox

NMDA receptors cloned at last

Mark L. Mayer

THE NMDA subtype of glutamate receptor, the holy grail of neurotransmitter receptor molecular neurobiology, has at last been cloned. Involved in excitatory synaptic transmission, learning, the regulation of neuronal growth, as well as ischaemia and epilepsy, the NMDA (*N*-methyl-D-aspartate) receptor ion channel complex has become the focus of a multimillion-dollar research effort for both the academic neuroscience community and the pharmaceutical industry. On page 31 of this issue, a paper from Nakanishi's laboratory describes the isolation by expression cloning of a functional NMDA receptor subunit¹.

The various receptors in the central nervous system for the excitatory amino-acid neurotransmitter glutamate fall into two main classes. One comprises the ligand-gated ion channels exemplified by the NMDA and the AMPA/kainate (non-NMDA) receptors; the other comprises the metabotropic receptors, which are coupled with G proteins. A functional AMPA/kainate receptor subunit was first cloned a few years ago². Its nucleotide sequence indicates a large N-terminal extracellular region followed by four transmembrane segments. The NMDA receptor now cloned by Nakanishi's group is of similar size, possesses the large N-terminal portion and four membrane-spanning segments, and has sequence homology with the AMPA/kainate receptor subunits.

In a field filled with rumour and speculation, and in which advances come overnight, rapid progress in cloning AMPA/kainate receptors had contrasted with the dearth of success in isolating NMDA receptors. The unpublished negative results generated speculation that the NMDA receptor was likely to be assembled from several distinct subunits, and that, unlike the cloned AMPA/kainate receptor, homomeric receptors assembled from a single subunit would be inactive.

This would have been the kiss of death for the isolation strategy used by Nakanishi's group. Several lines of evidence supported such pessimism: attempts to purify the NMDA receptor using affinity chromatography³, as well as results from experiments using photoaffinity labelling⁴, suggested that the NMDA receptor channel complex was composed of several subunits, of relative molecular mass around 25,000–70,000 (25–70K). The discovery of multiple allosteric sites on NMDA receptors was also an indication that, as for GABA (γ -aminobutyric acid) receptors, several types of subunit might be required for assembly of fully

functional NMDA receptors. Nakanishi's group have now shown this not to be the case.

Also in this issue, on page 70, Michaelis and colleagues report the cloning of a different gene, from rat brain, encoding a glutamate-binding protein, which they propose as the glutamate-binding subunit of an NMDA receptor complex⁵.

The situation is reminiscent of an earlier issue of *Nature* in which three independent laboratories described the cloning of non-NMDA receptors and kainic acid-binding proteins^{2,6,7}. The functional NMDA receptor clone from Nakanishi's laboratory¹ brings to mind the isolation of an AMPA/kainate receptor subunit² by Heinemann's laboratory at the Salk Institute. In both cases, expression cloning isolated complementary DNAs coding for proteins of relative molecular mass around 100K which, when expressed by *Xenopus* oocytes, generated functional receptor channel complexes with the characteristic physiology and pharmacology of NMDA and non-NMDA receptors respectively, as characterized in intact tissue. In the case of the NMDA receptor now cloned by Nakanishi's laboratory, its activation by glutamate and glycine, together with blocking by selective NMDA receptor antagonists and the lack of activity of selective non-NMDA receptor ligands, strongly suggests that homomeric NMDA receptors behave in a remarkably similar manner to NMDA receptors in intact tissue. The first cloning of an AMPA/kainate receptor was rapidly followed by isolation of three families of non-NMDA receptors^{8–11}; the sequences of seven subunits have been published to date, with probably more to be discovered. As for non-NMDA receptors, as well as for GABA, glycine and nicotinic acetylcholine receptors, the report of Nakanishi and colleagues will undoubtedly stimulate the discovery of additional NMDA receptor subunits.

The glutamate-binding protein cloned by Michaelis and colleagues⁵ recalls earlier reports in *Nature* of kainic acid-binding proteins^{6,7}, whose relationship to non-NMDA receptors, as defined using physiological and autoradiographic techniques applied to intact central nervous system tissue, remains uncertain. In the case of the glutamate-binding protein, Michaelis and his colleagues used a cloning strategy based on an antibody raised against a glutamate-binding protein isolated from rat brain. The cDNA isolated encoded a protein of about 57K. Because little was known

about the molecular properties of NMDA receptors, the lack of ion channel function on expression of this cDNA at first glance seemed reasonable, and in agreement with results from NMDA receptor purification and photoaffinity labelling experiments. Despite this, one is struck by the parallels between these first attempts to study NMDA receptor molecular biology, and the earlier work on AMPA/kainate receptors. The irrefutable functional analysis performed by Nakanishi and his colleagues¹ leaves little doubt that they have succeeded in isolating a genuine NMDA receptor subunit. In contrast, one wonders if the glutamate-binding protein isolated by Michaelis could have some other function apart from a role in the formation of a receptor channel complex?

Some clues are available in helping to decide between these two models for the molecular composition of NMDA receptors. Radiation inactivation analysis¹² suggests that the protein(s) carrying the glutamate — and glycine — binding sites of NMDA receptors have relative molecular masses of approximately 120K, in good agreement with the size of the subunit cloned by Nakanishi's laboratory, but much larger than the size estimates obtained by attempts to purify or photoaffinity label the NMDA receptor channel complex.

But radiation inactivation analysis also suggests there is a 210K subunit required for binding of the competitive NMDA receptor antagonist CPP, an antagonist used by Nakanishi to characterize functional activity of the cloned NMDA receptor. Molecular modelling by Ortwin, Bigge and colleagues resolves this difficulty. It suggests that NMDA receptor agonists and antagonists can indeed share a common recognition site, in which the α -amino, α -carboxyl, ω -carboxyl and ω -phosphonate groups in agonists and antagonists are similarly oriented for binding to the receptor. But these studies also suggest that a different recognition site on the NMDA receptor complex, perhaps attributable to the 210K subunit, can bind an additional oxygen atom in ω -phosphonate antagonists, thus accommodating both sets of results.

Quantitative analysis of the action of agonists and competitive antagonists on cloned NMDA receptors should reveal whether homomeric receptors are just similar to or functionally identical to NMDA receptors in intact cells. Experiments of this type are now being performed on AMPA/kainate receptors. They reveal marked differences in the kinetic behaviour of receptors assembled from cloned subunits from those of native receptors in hippocampal neurons, suggesting that quantitative analysis of receptor properties with Nobel-

approved patch-clamp techniques may provide clues to the subunit combinations required for assembly of naturally occurring glutamate receptors.

The NMDA and AMPA/kainate receptors cloned to date are strikingly different from those for other neurotransmitters, in that hydrophathy analysis suggests the presence of an unusually large N-terminal extracellular domain, such that glutamate-receptor ion channels have about twice the molecular mass of those for GABA, glycine and acetylcholine. A similar structural feature occurs in the metabotropic, G-protein-coupled glutamate receptors. Given this unique property, one might argue that isoforms could exist of glutamate receptors with lower molecular mass, similar to those for other transmitters, as suggested by the protein purification experiments on NMDA COSMOLOGY

receptors^{3,4}. The balance of evidence for NMDA receptors currently favours the 100K protein as the real McCoy, but further work could change this. With the polymerase chain reaction and lots of hard work the answers should be here soon. □

Mark Mayer is in the Section of Neurophysiology and Biophysics, NICHD, National Institutes for Health, Bethesda, Maryland 20892, USA.

1. Moriyoshi, K. *et al. Nature* **354**, 31–37 (1991).
2. Hollmann, M. *et al. Nature* **342**, 643–648 (1989).
3. Ly, A. M. & Michaelis, E. K. *Biochem.* **30**, 4307–4316 (1991).
4. Ikin, A. F., Kloog, Y. & Sokolovsky, M. *Biochem.* **29**, 2290–2295 (1990).
5. Kumar, K. N. *et al. Nature* **354**, 70–73 (1991).
6. Wada, K. *et al. Nature* **342**, 684–689 (1989).
7. Gregor, P. *et al. Nature* **342**, 689–692 (1989).
8. Keinänen, K. *et al. Science* **249**, 556–560 (1990).
9. Boulter, J. *et al. Science* **249**, 1033–1037 (1990).
10. Werner, P. *et al. Nature* **351**, 742–744 (1991).
11. Egebjerg, P. *et al. Nature* **351**, 745–748 (1991).
12. Honore T. *et al. Eur. J. Pharmac.* **172**, 239 (1989).

Quark soup: do not boil

M. Fukugita and C. J. Hogan

CALCULATIONS^{1,2} using a purpose-built computer show that as the Universe cooled immediately after the Big Bang, the initial plasma of unbound quarks and gluons condensed smoothly into the more familiar subnuclear aggregates that we see now. The alternative, that the transition may have been discontinuous, would have had profound consequences for our picture of the subsequent evolution of the Universe.

For almost 20 years physicists have known the fundamental equations governing the strong, weak and electromagnetic interactions. As far as experiments have been able to determine, up to energies of the order of 100 gigaelectronvolts this standard model of matter and forces of nature³ is both correct and complete. In principle, quantum chromodynamics (QCD), as the theory of strong interactions is called, should allow us to compute the behaviour of matter at enormously high temperatures, above 10^{12} K — much higher than known to occur naturally in the present-day Universe. Although straightforward as a matter of principle, QCD calculations are difficult in practice, straining the most powerful computing resources. The new calculations bring us closer to the elusive goal of understanding the behaviour of matter at extreme temperatures not seen naturally since the Universe was ten microseconds old.

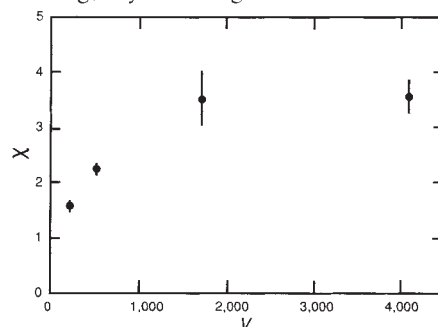
Analytical QCD predicts that, at very high temperatures, matter is converted into a quark–gluon plasma phase, sometimes affectionately called a quark soup. Matter in this phase consists not of protons and neutrons but of their liber-

ated constituent quarks. One of the most interesting questions is how suddenly the transition to this state occurs as matter is heated. For example, a sudden change in properties occurs as ice is converted to water or water to steam, and a similar sudden change in phase properties might occur as nucleons disintegrate into a gas of quarks. The issue is of practical interest in cosmology, as such an abrupt, first-order phase transition would cause a cooling, expanding Universe to boil or condense into macroscopic bubbles or droplets which can affect cosmic evolution, in particular the later production of nuclei whose abundances can be measured today.

Purely analytical attempts have failed to provide definite answers on the order of the physical transition. Persuasive arguments suggest that an idealized theory, with only extremely massive quarks — which is the same as having no physical quarks at all, but only gluonic force fields interacting amongst themselves — displays a first-order transition⁴. Also, a first-order transition occurs for three or more completely massless quark flavours in a closely related theory, the sigma model⁵. These model transitions differ in character; the first is characterized by the order parameter of 'confinement' and the second by that of 'chirality'. But for the true theory of nature, no proof exists of the behaviour; we cannot guess analytically whether real physical QCD, which has two light quarks (up, down), one heavyish one (strange), and three very heavy ones (charm, top, bottom), has a first-order transition.

Thus for the moment the only way to answer this question is by numerical simulations on a 'lattice', a discrete model of the continuous system of force fields and quarks described by QCD. Equilibrium values of physical quantities, such as order parameters characterizing the structure of the material in the model, can be computed as a function of temperature. The difficulty has been to know how much the dual approximations of discreteness and finite extent of the lattice introduce unphysical behaviour into the model system. The second problem is particularly acute as the lattices in numerical experiments are never very large; the number of operations is proportional to the number of links of a four-dimensional discrete lattice, which increases very rapidly with the number of elements.

With the rapid increase of computing power, the transition order has recently been examined using a powerful technique called finite-size scaling. There is no true singularity on any finite lattice and the study on a single size of lattice often leads to a misleading conclusion. Scaling, by showing how results vary



Lattice-volume (V) dependence of the susceptibility (mean-square fluctuations) of the chiral order parameter for finite-temperature QCD phase transition with two degenerate light quarks close to the physical mass: $m = 0.0125$ in lattice units for $V = 6^3, 8^3, 12^3$, and $m = 0.01$ for $V = 16^3$, the rightmost point, corresponding to $m \approx 7$ and 6 MeV, respectively, in physical units. A linear increase is expected if the transition is first order; this possibility appears to be ruled out by the data.

with changing size, provides the most decisive way known to statistical physicists for determining the order of phase transitions. Especially, it reveals the artefacts introduced by the finite extent of the lattice. This had not been done for QCD before because of the vast amount of computing time it requires.

A recent finite-size test by one of us (M.F.) and colleagues confirmed that the transition is first order for four flavours of light quarks⁶. For two flavours, there was no evidence of a first-order transition, but the lattice size was too small (space $\times$ time = $6^3 \times 4$ to $12^3 \times 4$) to rule it out either. Even so, these calculations took thousands of hours on

the fastest state-of-the-art conventional supercomputers.

Now A. Vaccarino *et al.*¹ from Columbia University report the results of a simulation on an even larger spatial lattice ($16^3 \times 4$), using a custom-made, 16-gigaflop superdupercomputer (a 16×16 grid of fast-array processors) run continuously for many months. These calculations were made with physical quark masses (5–10 MeV). Although the new result alone does not quite decisively rule out a first-order transition, such a conclusion can be drawn by combining it with the earlier result⁶ and applying a finite-size scaling test. The crucial evidence is that the susceptibilities of the order parameters do not increase from a 12^3 to a 16^3 lattice (see figure on previous page) whereas one would expect an increase proportional to the spatial volume — by $(16/12)^3$ or 2.3 times — if the transition is of first order.

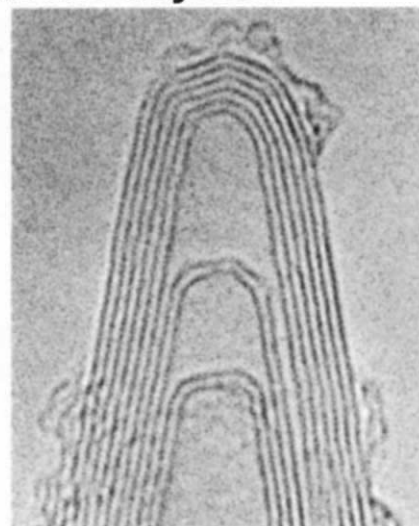
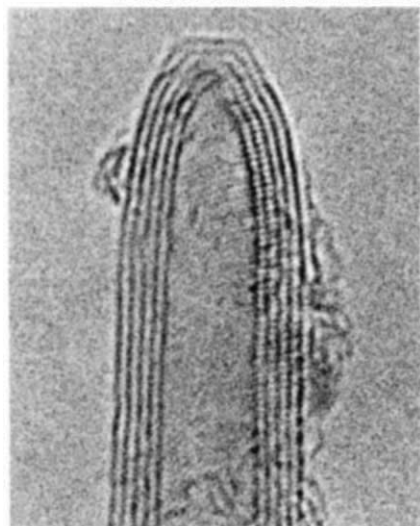
There now seems no possibility for even a mildly discontinuous, second-order phase transition for physical QCD. Although theoretically the transition in the two-flavour, zero-mass limit could be either of second or first order, a second-order transition is unstable against the perturbation which breaks the chiral symmetry responsible for the phase transition and it disappears as soon as the quark mass is turned on. Only a strong first-order transition would survive the perturbation of a finite quark mass.

Thus for just two flavours of physically realistic light masses corresponding to the up and down quarks there is no first-order transition. In principle, the strange quark (the next lightest after the up and down), were it light enough, could cause a first-order transition as the system approaches the behaviour of the sigma model. This is unlikely, as its mass (about 150 MeV) is too heavy to enhance the relevant chiral transition. In fact, the Columbia simulation with a strange quark put in the system showed that the situation remains unmodified, unless the mass of the strange quark is unphysically small².

The Universe progressed, it seems then, from a quark soup to the familiar hadron gas gradually; the nearly free quarks gradually congealed and eventually found themselves as they are today, completely confined in their hadronic bags. This means for example that

1. Vaccarino, A. *Nucl. Phys. B (Proc. Suppl.)* **20**, 263 (1991).
2. Brown, F. R. *et al. Phys. Rev. Lett.* **65**, 2491–2494 (1990).
3. Miller, D. J. *Nature* **349**, 279–287 (1991).
4. Yaffe, L. G. & Svetitsky, B. *Phys. Rev. D* **26**, 963–965 (1982).
5. Pisarski, R. D. & Wilczek, F. *Phys. Rev. D* **29**, 338–341 (1984).
6. Fukugita, M., Mino, H., Okawa, M. & Ukawa, A. *Phys. Rev. Lett.* **65**, 816–819 (1990); *Phys. Rev. D* **42**, 2936–2939 (1990).
7. Malaney, R. A. & Mathews, G. J. *Phys. Rep.* (in the press).

Needles in a carbon haystack



For those jaded by the sight of the C_{60} fullerene rolling across the pages of every other journal, the carbon structures described by Sumio Iijima on page 56 of this issue may provide a welcome change. He shows that the arc-discharge apparatus used for fullerene synthesis can, under certain conditions, produce needle-like tubules of graphitic carbon, up to a micrometre long and a few nanometres across, consisting of coaxial tubes nestling inside one another. These tubules reveal a new dimension to carbon's capacity for forming finite structures.

The needles grow on the cathode of the discharge apparatus with their axes aligned along the direction of the electric field, implicating the presence of the field in the formation mechanism. Following her studies of other forms of filamentary graphitic carbon, Mildred Dresselhaus has proposed previously that graphite sheets might curl up into tubes, but electron diffraction and microscopy of Iijima's structures indicate several surprising features. The coaxial tubes in a single needle adopt radii that allow a spacing between sheets almost identical to that in graphite, and the ends of

the needles are neatly sealed by caps which may be hemispherical, conical or faceted (as shown above). Therefore, like the fullerenes, the tubes are closed shells, albeit elongated with sturdy, laminated walls constructed from thousands of atoms.

Furthermore, each tube exhibits a helical twist, with rows of hexagons winding around the central axis. Iijima believes that this twist plays an essential part in the growth mechanism by ensuring that each successive carbon hexagon is added in a way that maintains regular axial growth, like adding steps to a spiral staircase.

When Iijima presented his results at the recent *Symposium on the Physics and Chemistry of Finite Systems* (Richmond, Virginia, 8–12 October 1991) there was no shortage of speculation about their implications. Could other atoms be trapped inside these tiny capillaries during formation, or could they be cracked open to pour such atoms inside? Might they be semiconducting, like graphite? The perfect crystallinity suggests high mechanical strength along the axis; needles grown to macroscopic lengths might constitute vastly superior carbon fibres.

Philip Ball

the standard model of nucleosynthesis in the early Universe, which neglects the effects of small-scale inhomogeneities, does not need modification for effects of QCD-induced bubbles⁷. (Of course, bubbles could still arise from even earlier events, but these tend to have a smaller spatial scale and more limited long-term effects.)

It is important to note that the results of these calculations do not apply to the material in neutron stars nor to high-energy heavy-ion collisions (currently being investigated at CERN and Brookhaven) in which QCD phase tran-

sitions may be sought. The early Universe, in which there were nearly equal numbers of particles and antiparticles in equilibrium with a hot bath of radiation, was an environment with an exceedingly high entropy. Realistic calculations for the colder, high-density environments accessible to observation, are beyond present computational capacity. □

M. Fukugita is in the Yukawa Institute for Theoretical Physics, Kyoto University, Kyoto 606, Japan. C. J. Hogan is in the Department of Physics, University of Washington, Seattle, Washington 98195, USA.

Face to face with babies

P. E. Bryant

MOST theories of perception used to depict young babies as perceptual incompetents, capable of discriminating little and of recognizing even less. Now psychologists are more likely to speak with awe than with contempt about the perceptual feats of neonates, and a paper by Johnson *et al.*¹, which appears in the latest issue of *Cognition* and deals with their attraction to human faces, serves as a striking reminder of how impressive their accomplishments are.

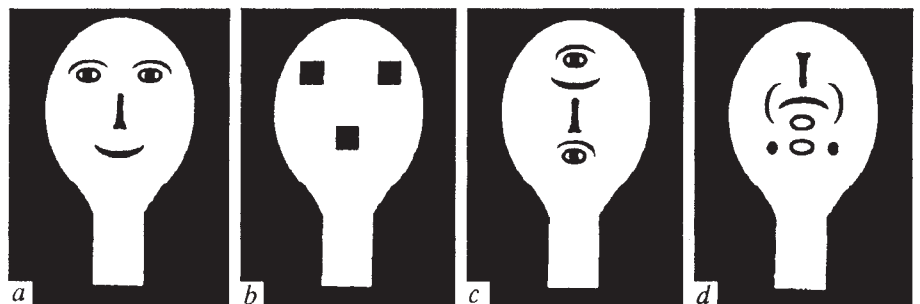
Nearly all recent studies of perception in young babies stem from the preference method invented by Robert Fantz². Fantz argued that, if you want to know whether infants see the difference between a pair of shapes, you should show them the two and if they consistently look at one for a longer time than at the other they must at some level have perceived the difference between the two. This method, and a subsequent refinement of it called the 'habituation paradigm', were a considerable success, and it is now clear that babies make many quite impressive discriminations from a very early age.

One, almost incidental, advantage of the preference method is that it also tells us a great deal about what babies like, as well as about what they can discern. The most riveting pattern for them seems to be a regular checkerboard. This dreadfully boring motif probably captivates the young infant because it contains a great deal in the way of contour and thus in the quantity of stimulation that it offers the infant.

Another, slightly less consistent, source of attraction for infants is the human face. In an ingenious, mobile version of the preference method, Goren and colleagues³ put various patterns, one at a time, in front of very young babies and then moved each pattern to one side: they reported that neonates followed the movement of a face-like pattern with more persistence than they did a pattern made from features of the face arranged in a random and definitely non-face like way (a 'scrambled face'). But it now seems evident⁴ that in their first month babies show no preference for looking at schematic faces rather than at 'scrambled' faces when these are presented statically, although two-month-old babies do.

The main aim of Johnson *et al.* was to explain why very young babies look longer at faces than at scrambled faces when the displays are moved, and yet apparently have no strong preferences of this sort when looking at static patterns. In two separate experiments, the authors

replicated the results of Goren *et al.*'s study. But the result by which Johnson *et al.* set most store, and which is their novel contribution, came from a third experiment which involved older babies who were one, three or five months old. This was also a tracking experiment, but it took a different form: this time the schematic face or the control patterns stayed in the same place and the babies were moved around the pattern in front of them. (The experimenters were forced to make this change because they found that the Goren procedure was not



Examples of stimuli used in experiments by Johnson and Morton. *a*, Face with all the right components in the right places. *b*, Face with the right arrangement but incorrect details. *c*, Linear face. *d*, Scrambled Face. (Taken from ref. 5.)

a success with the older infants.) Again, the question was whether babies would continue to look at the face longer than at the other patterns, and the answer — in many ways a surprising one — was that the one-month-old babies did but the older babies did not.

So, the story that Johnson and his colleagues tell is that babies of all ages prefer to look at human faces rather than at comparable patterns, but that the way that they show this preference is quite different at different ages. The authors' explanation for this developmental change is largely neurological. They appeal to the immaturity of the human visual system (particularly of the cortical visual pathways) at birth, and they argue that the preference for tracking mobile faces happens at a time when the babies' vision is largely under the control of subcortical mechanisms. Once the cortex takes over, the preference switches to a static one. This hypothesis is presented rather cursorily in the paper in *Cognition*, but is given a more expansive and comprehensible treatment in a new book⁵ by two of the paper's authors.

The hypothesis is a provocative one and raises many questions. One concerns the one novel result on which the hypothesis is based — the finding that one-month-old babies keep their eyes fixed on a face for a longer time than on any other comparable pattern when they

(the babies) are moved around the pattern, whereas older babies show no such preference. Although I can see why a more primitive visual system might function more effectively with a moving than with a static display, I cannot understand or find any explanation in the paper as to why older babies, which have more sophisticated mechanisms, should fail to track a face with particular persistence when they are moved around it. One mundane possibility, which is not part of the authors' scheme but does not seem to be ruled out by their data, is that the older children, now equipped with more powerful visual mechanisms, realize that the 'face' in question is not a real one and therefore do not bother to keep on looking at it when they are moved on.

Another difficulty, as far as I am concerned, is that there are other studies in which neonates have shown striking visual powers when confronted with static faces. One of these is a remarkable demonstration that neonates (only two days old) prefer to look at their mother rather than at another unknown woman⁶. So young babies' neurological mechanisms, however incomplete, do not prevent them from making some powerful discriminations about static human faces. Perhaps the restriction of preferences to moving displays in the early months applies only to schematic faces. We need more evidence on babies' preferences for real faces. Of course it is harder to provide the proper controls in a study with real faces: how do you scramble a real face? But such controls cannot be beyond the ingenuity which has always been the hallmark of research on infants' perception. □

P. E. Bryant is in the Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK.

1. Johnson, M. H., Dziurawiec, S., Ellis, H. & Morton, J. *Cognition* **40**, 1–21 (1991).
2. Fantz, R. L. *Sci. Am.* **204**, 66–72 (1961).
3. Goren, C. C., Sarty, M. & Wu, P. Y. K. *Pediatrics* **56**, 544–549 (1975).
4. Maurer, D. & Barrera, M. *Child Development* **52**, 196–292 (1981).
5. Johnson, M. H. & Morton, J. *Biology and Cognitive Development: The Case of Face Recognition* (Basil Blackwell, Oxford, 1991).
6. Bushnell, J. W. R., Sai, F. & Mullin, J. T. *Br. J. dev. Psychol.* **7**, 3–16 (1989).

Astronomy versus astrophysics

David Lindley

GAMMA-ray bursts might have been designed expressly to puzzle astronomers. First detected about 20 years ago, their name says just about everything there is to say about them. They are bursts of γ -rays, lasting anything from tens of milliseconds to hundreds of seconds, and except for weak X-ray emission in a few cases they produce no identifiable radiation in any other waveband. Several hundred have now been catalogued. They are seen all over the sky and, with no intrinsic measure of their distance available, it has been impossible to say whether they are modest nearby events or hugely energetic phenomena at cosmological distances. A growing consensus that they originate in our Galaxy was upset last month by the first results from the Gamma-Ray Observatory (GRO), launched by the space shuttle last year. Its survey of about 100 bursts shows almost perfect isotropy in its distribution across the sky, not the flattened distribution expected of the disk of our Galaxy. A recent conference* saw about a hundred participants (one for every burst) trying, but generally failing, to sort out the contradictions.

Before the GRO announcement, there were indirect intimations of a galactic distribution. T. Murakami *et al.* (*Nature* **335**, 234–235; 1991) in 'snapshot' spectra of a γ -ray burst caught by the Ginga satellite, saw features at about 20 and 40 keV that they interpreted in terms of the cyclotron spectra of electrons spiralling through magnetic fields. The field they deduced is of the order of 10^{12} gauss, typical of a strongly magnetized neutron star.

More recently, J. L. Atteia *et al.* (*Nature* **351**, 296–298; 1991) claimed to have identified, in the distribution of 250 bursts collected by three Soviet–French satellites, anisotropy consistent with a population of sources confined to the galactic disk. But the GRO results (C. A. Meegan *et al.* *IAU Circ.* No. 5358; 1991) produced the opposite conclusion: its collection of 117 bursts is distributed quite isotropically.

Equally telling was the value of something called $\langle V/V_{\max} \rangle$, where the brackets indicate an average over all bursts. Although the distance of the bursts is unknown, it is possible to learn something about the distance distribution on the assumption that brightness is, on average, a proxy for distance. If the (unknown) distance of a burst is D , and the greatest distance at which a burst of that intrinsic brightness could be seen is

$D_{\max}$, the quantity $D/D_{\max}$ is simply related to the observed brightness of the burst above detector threshold. $D/D_{\max}$ is then cubed to produce $V/V_{\max}$, which is the volume contained by a sphere that just reaches out to the burst divided by the maximum volume the burst could have occupied and still be seen. This is small for bright objects, and reaches unity for a burst that is barely detected.

For a distribution of equally bright objects reaching to infinity, $\langle V/V_{\max} \rangle$ is 0.5; but if the distribution goes out only to some finite distance within the detector's reach, $\langle V/V_{\max} \rangle$ decreases because there is then a paucity of faint objects. The 117 GRO objects yielded $\langle V/V_{\max} \rangle = 0.34 \pm 0.03$. Along with the isotropy this seems to imply an unlikely distribution, spherical around the Earth's position but also finite in radial extent. The let-out, however, is that a cosmological distribution of sources, although infinite, will also give a value of $\langle V/V_{\max} \rangle$ less than 0.5, because redshift reduces the visibility of distant objects.

Atteia *et al.* found $\langle V/V_{\max} \rangle = 0.42$; the Ginga survey produced a value of 0.35 (A. Yoshida, Riken Institute, Japan); bursts detected by the Solar Maximum Mission (SMM; G. Share, Naval Research Laboratory, Washington) and by various Soviet satellites (I. Mitrofanov, Space Research Institute, Moscow) likewise produced values of $\langle V/V_{\max} \rangle$ significantly less than one half. The only exception to this litany, pointed out by E. Fenimore (Los Alamos), came from a γ -ray detector aboard the Pioneer–Venus Orbiter (PVO), a satellite that was launched in 1978 and has been circling Venus since then. Its catalogue of more than 100 bursts gives $\langle V/V_{\max} \rangle = 0.46 \pm 0.04$, the only estimate consistent with 0.5.

A good way to compare these results is to think of the typical timescale on which individual experiments see bursts. GRO is finding one burst a day, whereas PVO sees about 10 a year. GRO will have seen by now perhaps one burst of the kind typically seen by PVO, and if PVO sees only the brightest, so closest bursts, it will plausibly come up with $\langle V/V_{\max} \rangle$ close to one half. On the other hand, SMM also compiled a collection of 100 bursts in about a decade, so that it and PVO would seem to be sampling the same population — but they produce different results.

The other news from GRO was that, so far, its spectrometers have seen nothing resembling the supposed cyclotron lines (B. Schaefer, Goddard Space Flight

Center), whereas Ginga has by now found three cases in 23 bursts bright enough to produce useful spectra.

Does this all make sense? The simple answer is no. The presumed galactic distribution and the presence of cyclotron lines in a few spectra has been the impetus behind some persuasive work on the generation of γ -ray bursts by neutron stars (A. Harding, Goddard Space Flight Center; D. Lamb, University of Chicago). To produce a continuum of γ -ray wavelengths, a favoured model is upscattering of soft X-rays to higher energies by fast electrons and positrons. The cyclotron features must then be imposed by the passage of γ -rays through a thin magnetized plasma.

It is not clear whether the explanations of the continuum and lines are compatible, or even internally consistent. For example, the 'Compton upscatter' models are for the most part based on an approximation in which a soft X-ray gets booted upwards in energy once only. But J. Brainerd (Marshall Space Flight Center) argued that if conditions allow scattering at all, there will be multiple scattering. The danger is that if the medium is too opaque and the intrinsic energy of the bursts too large, the whole thing will blow apart: bursts must be no more than a few hundred parsecs away for this model to work.

Nevertheless, as the neutron-star model is the only one developed, its deficiencies are understandable. But the new GRO results require its adherents to believe that there is a large population of neutron stars associated with the Galaxy, but distributed more isotropically than any other component. As B. Paczynski (Princeton University) put it, the astrophysics of γ -ray bursts is at odds with the astronomy. Neutron stars in the galactic disk fail the isotropy tests, and putting a population of neutron stars into the galactic halo is not easy: the core radius of such an extended population would have to be tens of kiloparsecs at a minimum, but the radii of dark halos around other galaxies, deduced from rotation curves, are typically less than 10 kiloparsecs. R. Lingenfelter (University of California, San Diego) offered a two-component model in which the low $\langle V/V_{\max} \rangle$ is attributed to the combination of a faint local population responsible for those bursts with cyclotron lines and soft X-ray emission, and an intrinsically brighter population, perhaps high-velocity neutron stars propelled into the galactic halo, which contributes all the low $V/V_{\max}$ scores.

In the circumstances, according to Paczynski, the conservative choice must be that bursts are cosmological in origin. T. Piran (Centre for Astrophysics, Cambridge) proposed the coalescence of two neutron stars as the source. The avail-

*Gamma-ray Burst Workshop, Huntsville, 16–18 October, 1991.

able energy, comparable to a supernova output, is easily enough to power cosmologically visible bursts, and the coalescence of one pair per galaxy per 10^8 years is enough to explain the observed burst rate. In our own Galaxy there are at least two cases where pulsars are spiralling into neutron-star companions, and coalescence is the inevitable endpoint.

But coalescing neutron stars would release their energy mostly in the form of neutrinos (as does a supernova). Piran's idea is that the outburst is dense enough for neutrino pairs to generate electron-positron pairs that produce γ -photons. But the difficulty then is in preventing the resulting fireball from

becoming optically thick, and creating everything from radio waves up. Paczynski suggested that a neutron star being swallowed by a black hole might create a cleaner event.

The good news is that GRO collected two more bursts during the meeting, and is still finding one a day. The isotropy and $V/V_{\max}$ measurements must get better, and the quality of the spectra being taken is such that cyclotron lines will either be seen or will, with some certainty, be proved not to exist. In the meantime, as one participant said, it's time to "shut up and think". □

David Lindley is Associate Editor of *Nature*.

FLUID DYNAMICS

Forging the missing link

John D. S. Jones

IMAGINE a busy bar full of people smoking cigars and blowing smoke rings. Each ring is a cylinder of smoke swirling around its central axis and curled back on itself to form a closed loop. Clearly the development of a single smoke ring is a complicated phenomenon, but what happens when two collide? The internal motion within each smoke ring will significantly affect the final collision and it is difficult to predict the final outcome. For example, could two or more smoke rings eventually become linked? This is effectively the question H. Aref and I. Zawadzki answer on page 50 of this issue in their computer simulations of colliding vortex rings.

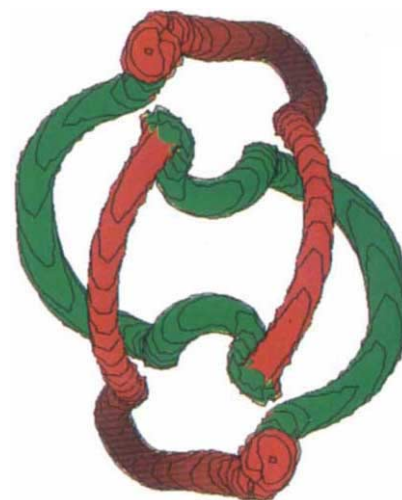
We are in the realm of fluid dynamics, that is, the study of the way fluids (or gases) flow. The vorticity of a fluid flow is a measure of the amount of local rotation in the flow. In two dimensions, there is vorticity when, for example, there is a fixed point in the plane around which nearby points tend to rotate. A specific example is flow in which all the points circulate around a single, fixed point in closed orbits. Another is when the points of the plane spiral into a central, fixed point (here, of course, there are no closed orbits).

In three dimensions, the simplest example of vorticity is simple rotational or spiral flow in a plane which may be moving along the third dimension. A vortex tube is, as with the smoke ring, given by starting with such a flow with vorticity in a cylindrical tube, for example simple rotational or spiral flow, deforming the tube by twisting and even knotting it. Finally, combine the internal flow in the tube with some time evolution of the tube itself. It is not difficult to imagine situations in which vortex tubes occur, for example in tornadoes and water flowing out of a bath.

A vortex ring occurs when the two ends of the vortex tube join up. The geometry and topology of vortex tubes and vortex rings are part of the domain of topological fluid dynamics. Aref and Zawadzki are interested in the dynamics and interactions of vortex rings, that is the way vortex rings evolve and what happens after they collide. Of course the mechanism would be simple if we were dealing with rings that contain no fluid — as the rings collide, one passes through the other and we achieve a state with linked rings. The question is really whether this can be done with the fluid swirling around inside the tubular rings and what happens to the fluid in this process.

This brings us to an important point — the viscosity of a fluid, which is a measure of the tendency of individual particles in the fluid to stick together, or, to put it another way, the internal friction in the fluid. If the fluid is perfect — has no viscosity — unlinked vortex rings cannot become linked. This observation goes back to Lord Kelvin and was instrumental in his 'vortex theory of atoms', developed in collaboration with P. G. Tait. The use of knots and links led Tait to attempt to classify knots in increasing order of complexity (P. G. Tait, *Scientific Papers I*, Cambridge University Press, 1898) and even though the vortex theory of atoms turned out to be misconceived, the catalogue of knots Tait produced and several of the conjectures he was led to are still important in the thriving modern theory of knots.

The equations governing the motion of fluids, the Euler equations for perfect fluids and the Navier-Stokes equations for viscous fluids, are complicated non-linear equations and it is really quite impractical to attempt to solve them analytically in terms of closed formulas.



Mix and match: two linked vortex rings, from one of Aref and Zawadzki's simulations. The colours identify the original rings.

However, there is an extensive literature on numerical, computational solutions to these equations. Using these methods, Aref and Zawadzki produce their examples of flows in which unlinked vortex rings collide and end up linked. The best way to get a feel for the mechanisms involved in these interactions is to look at the pictures of the flows Aref and Zawadzki give in their paper. Two intriguing features emerge. First, the authors find it easier to run the simulations backwards, starting with linked rings, to find initial conditions that lead to linking. Second, in the process of linking, the two rings lose their initial identity and become hybrids of one another (see figure).

Vortex tubes and vortex rings occur in many examples of fluid flows: for example, an aircraft taking off creates a vortex tube at the trailing edge of its wing. It is clear that the study of their evolution and interactions will be of some importance in both the qualitative and quantitative understanding of flows.

It is interesting that although these vortex interactions are essentially geometrical phenomena, it requires massive computing power to construct the simulations. Indeed, geometry is one of the few areas of mathematics that have been genuinely influenced by the use of computers.

Three questions arise. First, is it possible to describe this phenomenon of vortex rings colliding and linking in purely geometrical terms? Second, is it possible to characterize some good set of initial states that lead to linking? Lastly, can this be done experimentally — perhaps with smoke rings? □

John D. S. Jones is in the Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK.

Soft core

To test whether your egg is hard boiled or uncooked, spin it and see what happens. J. L. Hilton proposes something similar for Mars to see if it has a solid core (*Astr. J.* **102**, 1510–1527; 1991). Disturb a spinning body and its axis will continue to oscillate, or 'nutate', the Earth's 'Chandler wobble' being an example of this effect. But the exact response depends on whether the interior is rigid or not. Earth's molten core is just discernible in this way. Hilton calculates the effect of the torques exerted on a completely solid Mars by the Sun, Jupiter, its moons Phobos and Deimos, and its varying orbit; the effects of a liquid core are to be added in a subsequent paper. The test for a liquid core could, in principle, be applied to any planet, but only Mars gives us the clear view necessary to see its tiny nutation of 7 arcseconds a year.

Not in the blood

THE powerful disruptive biochemical effects of ozone, the three-atom allotrope of oxygen, may come to the assistance of those seeking to prevent transmission of the human immunodeficiency virus through blood products. K. H. Wells *et al.* report in *Blood* (**78**, 1882–1890; 1991) that the viral activity of HIV-infected blood samples treated with 1,200 parts per million of ozone is reduced 100-billionfold *in vitro*, without seriously affecting the plasma component factor VIII. Why this happens is unclear; perhaps the ozone attacks the fatty viral envelope and destroys the virus's capacity to infect cells, the authors speculate. Whether the treatment will make a feasible alternative to screening remains to be seen. Tests of the infectivity of treated blood in sensitive animal models will be an essential next step.

Malignant diagnosis

THE last thing one needs during an operation is a serious reaction to the general anaesthetic, but for sufferers of the rare disorder malignant hyperthermia (MH), the resulting sudden paralysis can be fatal if unchecked. Linkage studies have incriminated the ryanodine receptor (RYR1), which controls release of calcium ions in muscle cells, but a putative mutation has only just been identified. E. F. Gillard *et al.* (*Genomics* **11**, 751–755; 1991) describe a point mutation in the RYR1 gene in affected members of one out of 35 MH families, which substitutes an arginine residue for cysteine. The analogous mutation is found in five strains of pig that suffer the related porcine stress syndrome. However, the molecular basis of most MH cases remains to be determined, and may not be confined to the RYR gene.

Molecular metamorphosis

Dagmar Ringe and Gregory A. Petsko

NEVER throw away your wide ties, for one day they are certain to come back in style. This is a useful tip in science as well — today's backwater is often tomorrow's hot field, the proteases being a good example. Enzymology began with these digestive enzymes, but by the 1970s they were considered *passé*. Now they are centre-stage once more, figuring prominently in the design of drugs against AIDS, emphysema, arthritis and hypertension. On page 37 of this issue¹, a new Achilles' heel is presented for antiviral agents: the three-dimensional structure is described of a viral protease that in effect shoots itself in the foot so that it can assume a new identity as a structural protein.

Because they seem to come in a small number of families, with well-defined structures and modes of action, proteases are everyone's favourite molecules for model building from sequence similarity. Proteases act as cutting tools to tailor proteins and their precursors to fit their context in the execution of cellular processes. From the three-dimensional structure of the core protein of Sindbis virus, an enveloped virus of the alphavirus family, it now seems that some proteases are also capable of metamorphosis¹. Choi *et al.* find that the principal domain of the core protein has a polypeptide fold similar to that of the serine protease chymotrypsin. It possesses a complete catalytic triad, but its own C-terminal amino acid, a tryptophan, is lodged in its primary substrate specificity site, rendering it inactive. Thus, Sindbis-virus core protein is an enzyme designed to carry out a single turnover, cleaving itself from the viral polyprotein precursor and simultaneously changing itself into a structural protein.

Most single-stranded RNA viruses code for proteases of the chymotrypsin family, although some of these enzymes have cysteine replacing serine as the active site nucleophile. (Human immunodeficiency virus protease is a notable exception — it belongs to the structurally and chemically distinct family of aspartyl proteases.) These enzymes cleave the individual viral proteins from the giant polyprotein precursor in the first step of viral assembly. But they are not incorporated into the finished virus particle, and their function can sometimes be replaced by endogenous cellular proteases.

It has long been suspected that Sindbis virus and other members of the insect-transmitted alphavirus family are different. Simmons and Strauss showed in 1974 that the virus core protein, which

forms the coat surrounding the RNA, is responsible for its own release from the polyprotein². Sequence analysis³ and mutagenesis experiments^{4–6} indicated that Sindbis core protein has a catalytic triad of Asp 163, His 141 and Ser 215, similar to the Asp/Glu-His-Ser triad found in all serine proteases and in many esterases^{7,8}.

But this active-site similarity did not in itself establish that the three-dimensional structure of the Sindbis core protein is similar to that of, say, chymotrypsin, because many different polypeptide chain folds (for example, that of subtilisin) possess this same catalytic triad as a result of convergent evolution to a common function⁹. Moreover, the usual polypeptide chain fold for viral coat proteins is an eight-stranded 'jellyroll' antiparallel β -barrel which does not resemble any serine protease¹⁰. Indeed, it has been suggested, on the basis of a low level of sequence similarity, that Sindbis core protein should adopt the jellyroll fold, with a catalytic triad arising by convergent evolution¹¹. This prediction seemed reasonable as no structural protein was known to be folded like a serine protease.

The three-dimensional structure of the Sindbis core protein was determined by Choi and co-workers by X-ray crystallography, which establishes that the molecule possesses the chymotrypsin fold¹. The catalytic triad is in the same position as in the mammalian or bacterial chymotrypsin-like enzymes. There are, however, some significant and intriguing differences. There are no disulphide bridges. The 'methionine loop' is missing and so is the C-terminal helix; as already mentioned, the C terminus is folded back into the active site. The 'second serine' found in all other chymotrypsin-like proteases, which hydrogen-bonds to the side chain of the catalytic aspartate, is a non-hydrogen-bonding leucine in Sindbis core protein. The loops that form the P2 and P3 substrate-binding pockets in other serine proteases are completely absent. And finally, the catalytic aspartate 163, whose counterpart is buried in all other serine proteases, is exposed to solvent.

Ordinarily, deployment of a protease as a structural protein would be like using termites to construct a hardwood floor. Sindbis virus solves this problem by having the core protein inactivate itself. The *cis* and *trans* cleavage of a protein from a viral precursor cannot normally be distinguished, but in the case of Sindbis core protein the X-ray structure shows that autocatalytic scis-

sion must have occurred by *cis* cleavage: no other mechanism would leave the protein's own C-terminal residue in the active site after processing.

Evidence is mounting that many proteases are their own best substrates. Pepsin is one example: the propeptide obstructs the active site before undergoing autocatalytic processing¹². γ -Chymotrypsin as isolated contains a bound peptide that is probably derived from auto/lysis¹³. Sindbis-virus core protein is shown here to be another example. Given that the function of a protease is, like any enzyme, to bind substrate and release product to make way for another turnover, the Sindbis core-protein structure raises the obvious question of why the C-terminal segment remains bound in the active site and why, even then, other substrates cannot displace it and be hydrolysed.

Maybe they can. There is not much known about the enzymology of this protease, although that should change rapidly. But, assuming the mature molecule really is inactive, we can speculate on how that happens. One possibility is that the C-terminal peptide is bound so tightly that no substrate can displace it. That might seem unlikely, given the absence of interaction at the P2 and P3 subsites, but it is still possible because this inhibitory peptide is a part of the covalent and folded structure of the protein. The entropic cost of immobilizing it in the active site has already been paid by the free energy of folding.

Another possibility is that the catalytic machinery is defective. Asp 163, the anionic member of the catalytic triad, is exposed to solvent in this structure, which should have a substantial influence on the pK_a of the catalytic histidine to which it is hydrogen-bonded. Certain pyrone derivatives inactivate chymotrypsin by, in part, altering the histidine pK_a through interaction with a solvent-exposed carboxylate¹⁴. Presumably, when Sindbis-virus core protein is ini-

tially folded as part of the precursor polypeptide, the C-terminal residues that are ultimately cleaved off shield Asp 163 from solvent before they dissociate, allowing the first, and only, turnover event.

Finally, it is possible that the bound C terminus is not a free carboxylic acid but is actually a covalent inhibitor attached as an acyl group to the catalytic serine 215. At 3-Å resolution, Choi *et al.* cannot be certain about this. There is precedence: high-resolution X-ray diffraction shows that the inhibitory peptide in γ -chymotrypsin seems to form a stable acyl enzyme in just this fashion¹³. Crystallographic studies of inhibitors bound to chymotrypsin suggest that the structural and chemical requirements for deacyla-

EVOLUTION

Sex, slime and selfish genes

Laurence D. Hurst

THE initial evolution of sex, meaning the transfer of genetic information and the fusion of gametes, has attracted much less interest than the reasons for maintaining sex. But work by Meland *et al.*¹ and Kawano *et al.*² on the genetics of a slime mould provides empirical support for one of the most original speculations on the origin of sex, as well as unsettling some old evolutionary truisms.

The life cycle of *Physarum polycephalum* includes two distinct vegetative forms: a uninucleate amoeba and a multinucleate jelly, termed the plasmodium



Photo courtesy of R. Anderson.

"Yea, slimy things did crawl...": a single plasmodium of *Physarum polycephalum*, which has grown to cover a petri dish.

(see figure). The plasmodium can cover many square centimetres and contain millions of diploid nuclei which divide synchronously. As the jelly moves it leaves behind the protective extracellular slime sheath that gives the organism its unbecoming title. Meiotic division of the diploid nuclei results in the amoebae which act as gametes. These fuse in pairs

to form diploid zygotes which in turn differentiate into the plasmodium.

Unlike the sperm and eggs of many organisms, all of the amoebae are the same size (isogamous). But despite the fact that both gametes have their own mitochondria, only one set of mitochondrial genes is found in the plasmodium^{3,4}. Meland *et al.*¹ show that this pattern of inheritance involves an active mechanism involving the degradation of one of the sets of mitochondria. The authors point out that this may be the first report of the active degradation of mitochondria in sexual crosses, although the possibility has been considered previously⁵.

Similar mechanisms have been proposed to explain the uniparental inheritance of chloroplasts in isogamous single-celled algae such as *Chlamydomonas reinhardtii*⁶. *Chlamydomonas* gametes are of one of two types: plus or minus. Plus will mate only with minus, and vice versa. Following fusion, the minus-type chloroplast DNA is degraded⁶. Some time ago, Grell⁷ remarked that as regards the sexual differentiation of isogametes which fuse to form a zygote, all instances were of the *Chlamydomonas* type and that "no case of more than two sexes has ever been found". Following the work of Meland *et al.*¹ and Kawano *et al.*⁴ on *Physarum*, this statement might need revising. Laboratory studies indicate that any *Physarum* gamete can mate with any other so long as they are genetically non-identical at all of three different polymorphic loci⁸. The inheritance of mitochondria is determined by what appears to be a linear hierarchy of at least 13 alleles at one of these three loci. For instance, knowing that mitochondria from

1. Choi, H. K. *et al.* *Nature* **354**, 37–43 (1991).
2. Simmons, D. T. & Strauss, J. H. *J. molec. Biol.* **86**, 397–409 (1974).
3. Boege, U. *et al.* *Virology* **113**, 293–303 (1981).
4. Hahn, C. S., Strauss, E. G. & Strauss, J. H. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4648–4652 (1985).
5. Melancon, P. & Garoff, H. *J. Virol.* **61**, 1301–1309 (1987).
6. Hahn, C. S. & Strauss, J. H. *J. Virol.* **64**, 3069–3073 (1990).
7. Winkler, F. K., D'Arcy, A. & Hunziker, W. *Nature* **343**, 771–774 (1990).
8. Sussman, J. L. *et al.* *Science* **253**, 872–879 (1991).
9. Wright, C. S., Alden, R. A. & Kraut, J. *Nature* **221**, 235–242 (1969).
10. Rossman, M. G. & Johnson, J. E. *A. Rev. Biochem.* **58**, 533–573 (1989).
11. Fuller, S. D. & Argos, P. *EMBO J.* **6**, 1099–1105 (1987).
12. Sielecki, A. R., Fujinaga, M., Read, R. J. & James, M. N. G. *J. molec. Biol.* **219**, 671–692 (1991).
13. Dixon, M. M. & Matthews, B. W. *Biochemistry* **28**, 7033–7038 (1989).
14. Ringe, D., Mottonen, J. M., Gelb, M. H. & Abeles, R. H. *Biochemistry* **25**, 5633–5638 (1986).
15. Stoddard, G. L., Bruhnke, J., Porter, N., Ringe, D. & Petsko, G. A. *Biochemistry* **29**, 4871–4879 (1990).

type α parents expunge type β , and that type β wins over type γ , it follows that type α will win over type γ (refs 1, 4).

Mitochondria in *Physarum* are however not always inherited uniparentally. Kawano *et al.*² show that the above is all true if the mitochondria from both parents lack a particular plasmid. But if one of the mitochondrial sets has the plasmid (the *mif*⁺ condition), whereas the other does not (*mif*⁻), then the mitochondria fuse, undergo recombination and subsequently divide within the zygote. Most or all of the derived mitochondria inherit the plasmid in an unaltered form. When two *mif*⁻ gametes mate mitochondrial fusion is not witnessed, indicating that organelle fusion is a direct consequence of the possession of the plasmid. The consequences of fusion between two *mif*⁺ gametes is unknown. As Kawano *et al.* comment, the behaviour of mitochondria in *mif*⁺/*mif*⁻ matings appears to constitute some sort of mitochondrial meiotic cycle.

More significantly, a system such as this is very similar to that proposed by Hickey and Rose⁹ for the initial stage in the evolution of sex. Attempting to understand the origins of both the transfer of genes from one individual to another and of the fusion of gametes, Hickey and Rose proposed that a parasitic gene might underlie the process. Consider a transposon or similar selfish gene in a cell. If this gene is incapable of horizontal transmission, it is stuck in the vertical lineage of its host cell. But what if this gene can force its cell to fuse with another, leave a copy of itself in this cell, and then force the two cells to split apart? The gene would receive a two-fold transmission advantage while enduring little cost. Hence we might expect such a gene to spread. Kawano *et al.*'s observation of the mitochondrial plasmid in *Physarum* is direct support for the possibility that selfish genes are capable of manipulating their hosts in the manner Hickey and Rose propose.

Recombination between mitochondrial genomes has also been found in a few other organisms, the best studied of which is yeast¹⁰. Recombination can involve biased gene conversion, or a similar process at the ω locus such that in a cross between cells which are ω ⁺ and ω ⁻ nearly all of the mitochondria end up being ω ⁺ (ref. 10). Unlike the *Physarum* case, ω ⁺ is not a plasmid but an intron¹⁰. No biased conversion is seen in crossing of cells with like-type mitochondria. But it is not known if ω ⁺ can induce mitochondrial fusion. Recombination is known between mitochondria in the absence of ω ⁺. The most commonly cited case in support of Hickey and Rose is that of the F plasmid of the bacteria *Escherichia coli* which is responsible for the production of a repro-

ductive pilus through which a copy of the plasmid can migrate from one cell to another¹¹. In this system, fusion of bacteria is not seen and the recombination is found only in the recipient bacterium.

Could a selfish genetic element be responsible for fusion between eukaryotic cells? Once again, evidence comes from *Physarum*. One of three consequences ensue when two plasmodia meet¹². If they are genetically identical at at least four particular loci, then the two plasmodia fuse; this fusion is a vegetative process not involving nuclear fusion. If the plasmodia differ genetically then either no fusion occurs, or fusion ensues followed by a post-fusion incompatibility reaction. This reaction involves the nuclei of the partner bearing a particular recessive allele being condensed, enclosed in vacuoles and eliminated from the cytoplasm. The other nuclei containing the dominant *killer* gene are unaffected. Hence, as with the ω locus, the F plasmid, *Physarum*'s mitochondrial plasmid and similar selfish genes, the aggressive dominant gene increases in frequency at least in the plasmodium. The fate of the mitochondria after fusion of the two plasmodia is unknown.

All this is consistent with a Hickey and Rose type view of the evolution of fusion. But what of the evolution of recombination? The independent segregation of chromosomes at meiosis follows as a consequence of the fusion of haploid gametic nuclei. Why, in most organisms, do these nuclei not remain as two independent entities as they do in some fungi? The post-fusion incompatibility reaction seen in *Physarum* also involves the induction of nuclear fusions. Lane and Carlile¹² suggest that by fusing with a nucleus with the dominant gene, the nuclei with the recessive gene might escape destruction. Could the fusion of haploid nuclei be a result of selection to avoid the deleterious consequences of similar competition between the two genetically different nuclei? □

Laurence D. Hurst is in the Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.

- Meland, S. *et al.* *Curr. Genet.* **19**, 55–60 (1991).
- Kawano, S. *et al.* *Protoplasma* **160**, 167–169 (1991).
- Kawano, S., Anderson, R. W., Nanba, T. & Kuroiwa, T. *J. gen. Microbiol.* **133**, 3175–3182 (1987).
- Kawano, S. & Kuroiwa, T. *J. gen. Microbiol.* **135**, 1559–1566 (1989).
- Eberhard, W. G. *Q. Rev. Biol.* **55**, 231–249 (1980).
- Kuroiwa, T. *Microbiol. Sci.* **2**, 267–270 (1985).
- Grell, K. G. *Protozoology* 225 (Springer, Berlin, 1973).
- Kawano, S., Kuroiwa, T. & Anderson, R. W. *J. gen. Microbiol.* **133**, 2539–2546 (1987).
- Hickey, D. H. & Rose, M. R. in *The Evolution of Sex* (eds Michod, R. E. & Levin, B. R.) 161–175 (Sinauer, Sunderland, Massachusetts, 1988).
- Dujon, B. in *The Molecular Biology of the Yeast Saccharomyces* Vol. 1 (eds Strathern, J. N., Jones, E. W. & Broach, J. R.) 505–635 (Cold Spring Harbor Laboratory Press, 1981).
- Willems, N. S. & Skurray, R. A. *Rev. Genet.* **14**, 14–76 (1980).
- Lane, E. B. & Carlile, M. J. *Cell Sci.* **35**, 339–354 (1979).

Arresting waste

PLASTIC waste, says Daedalus, is not really immortal. Exposed to air and daylight, it gradually crumbles into small particles which are ultimately metabolized by bacteria. Only when buried in landfill dumps, out of reach of air and daylight, does it last indefinitely.

On this argument, plastic junk should simply be spread out on waste ground, or gathered into heaps, and left to weather. Daedalus now proposes to do just this, usefully. He points out that solvent vapours in the air can attack polymers, softening them and making their surface tacky. When the absorbed solvent evaporates again, the polymers recover their normal strength and character.

So DREADCO engineers are compacting domestic and commercial plastic waste into a sort of open-weave bonded structural material. The mass of detergent bottles, yogurt pots, plastic crockery and cutlery, and so on, are exposed to solvent-loaded air while passing between slightly compacting rollers. The items soften and stick together, and emerge from the drying chamber as an open but continuous material whose individual components, deformed and firmly adhering, are still discernible. The solvent vapour can be recovered for re-use. The waste need not be segregated into different polymers, or even freed completely from other waste. A few tin cans or cardboard boxes glued into the structure would not matter.

Some outdoor use has to be found for 'Wastebloc'® while it slowly breaks down. Daedalus proposes to use it as a central crash barrier for motorways. A vehicle ploughing into it would rapidly be slowed by the bending and shearing of thousands of plastic components, but would suffer little damage in return. In this way, the vast motorway network will usefully double as a waste tip. Wastebloc barriers, endless miles of deformed plastic refuse welded together into a thick monolithic fence, will of course be irredeemably ugly, thus harmonizing pleasantly with the motorway ambience. Plants and shrubs could usefully be encouraged to grow round and through the Wastebloc fencing. The resulting 'reinforced hedge' would be even more energy-absorbing; and Wastebloc will probably break down and biodegrade faster in the rich micro-ecology of a plant community. As fast as it weathers away, or is damaged by traffic, new plastic waste will be grafted onto it by special rubbish trucks fitted with compacting rollers and vapour-dispensing equipment. Thus plastic rubbish, the bane of modern waste disposal, will enjoy a long and useful retirement while nature slowly reclaims it.

David Jones

Hazard from volcanic ash

SIR — Volcanic ash clouds pose a significant hazard to aircraft in flight. There were at least 14 such incidents with the volcanic ash clouds arising from the recent eruptions of Mount Pinatubo in the Philippines. Commercial aircraft have no means of discriminating between volcanic ash clouds and normal water and ice clouds. We have developed a technique¹ that relies on passive infrared radiometry to detect and discriminate ash clouds. We have performed ground-based tests of a radiometer on volcanic eruption clouds from the Sakurajima volcano in southern Japan which confirm that the technique is viable.

Volcanic ash consists of glass and

CENTRAL WAVELENGTHS, BANDWIDTHS AND FUNCTION OF THE FIVE INTERFERENCE FILTERS USED IN THE PASSIVE INFRARED RADIOMETER

Wavelength (μm)	Bandwidth (μm)	Function
8.6	0.3	SO ₂ emission
10.1	0.5	Water/volcanic ash discrimination
10.9	1.1	
11.8	1.4	
12.0	0.6	

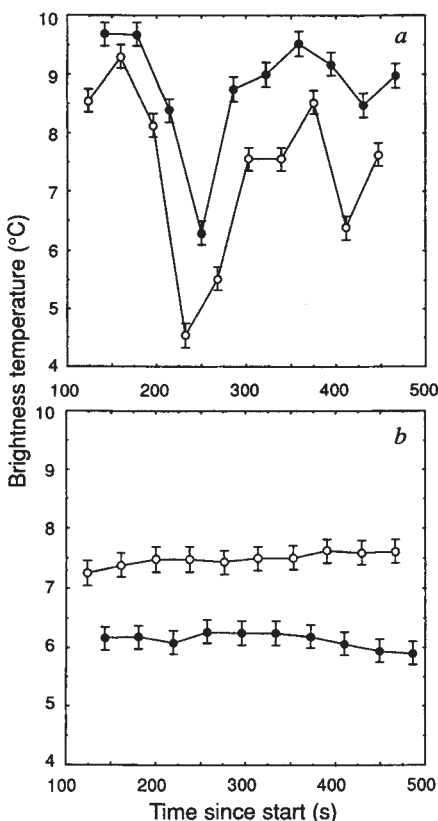
silicate-bearing minerals which melt under the high operating temperatures of jet engines and then fuse onto internal surfaces². The ingestion of fine particles is sufficient in some cases to limit the flow of air through the engines, causing them to stall. Other damage to the aircraft arises through abrasion of surfaces (in particular 'sand blasting' of external surfaces, including the cockpit windscreen), disruption of electronics and corrosion of external surfaces by sulphuric acid droplets.

The test instrument is a chopper radiometer using five interference filters (see table) to achieve spectral selection in the thermal infrared region of the electromagnetic spectrum, between wavelengths of 8 and 13 μm . During two weeks in March 1991, the radiometer was deployed at the Sakurajima Volcanological Observatory and at several locations on Sakurajima volcano (31° 35' N, 130° 40' E). Eruptions of the volcano take place between 100 and 200 times per year and between eruptions, the volcano continually emits large quantities of ash³. The radiometer was ground-based and could be moved in azimuth and elevation. The temperature measured by the radiometer (the brightness temperature) is obtained for selected filters after calibration. At one site, 3 km from the volcano, the radiometer viewed the ash-laden plume at an elevation angle of 23°. Here, only two filters were used. The measurements of the Sakurajima ash cloud indicate a unique signa-

ture — more radiation is received at 10.9 than at 11.8 μm (*a* in the figure), opposite to what is observed when viewing water clouds (*b* in the figure) and clear skies.

The difference between the brightness temperature at 10.9 and 11.8 μm (ΔT) is a sensitive function of the optical properties of the target cloud and the absorption and emission characteristics of atmospheric constituents between the cloud and the radiometer. For measurements at these wavelengths the principal absorber of radiation in the lower atmosphere is water vapour, which increases the radiation received at the longer wavelength. Thus for ground-based measurements, atmospheric water-vapour effects cause ΔT to decrease. The extinction of infrared radiation from clouds depends on the optical characteristics and thickness of the cloud, particle shapes and the distribution of particle sizes within the cloud⁴.

Radiative transfer calculations for a cloud of particles using available data on the refractive indices of quartz, water and ice indicate that the infrared extinction of quartz at 10.9 μm is greater than that at 11.8 μm , whereas for water and ice the extinction is less at 10.9 than at 11.8 μm (ref. 5). Thus ground-based and



a, Brightness temperature as a function of time for the 10.9 ($\bullet$) and 11.8 μm ($\circ$) filters, when the radiometer is viewing the volcanic plume. *b*, As *a* but the radiometer is viewing a water cloud at approximately the same elevation angle.

airborne measurements of ΔT should be positive for quartz and negative for water and ice clouds, in agreement with our measurements.

Even though the ground-based measurements were made with an intervening humid atmosphere, the volcanic ash signature was strong. At the altitudes used by commercial jet aircraft there is little water vapour and an airborne radiometer should easily detect the volcanic ash signature. Development of an airborne prototype is planned for full operational trials.

A.J. PRATA
I.J. BARTON

CSIRO Division of Atmospheric
Research,
Aspendale, Victoria 3195, Australia
R.W. JOHNSON

Bureau of Mineral Resources,
GPO Box 378, Canberra,
ACT 2601, Australia

K. KAMO
Sakurajima Volcanological Observatory,
Kyoto University, Sakurajima,
Kyushu, Japan

J. KINGWELL
CSIRO Office of Space Sciences &
Applications,
ANU, Acton, ACT 2601, Australia

1. CSIRO Int. Pat. No. PCT/AU91/00132 (1991).
2. Smith, W.S. *Nat. Hist.* **11**, 26–34 (1983).
3. Kamo, K. & Ishihara, K. *IAVCEI Proc. Volcanol* **1**, 585–598 (1989).
4. Jennings, S.G., Pinnick, R.G. & Auvermann, H.J. *Appl. Opt.* **17** (24), 3922–3929 (1978).
5. Prata, A.J. *Geophys. Res. Lett.* **16**(11), 1293 (1989).

Camouflaged DNA

SIR — I read with interest C. Weissmann's Review article (*Nature* **352**, 679–683; 1991) on the scrapie agent in which he introduces a new (holopriion) model to reconcile conflicting data and hypotheses in this highly controversial field. May I put forward a suggestion of my own? It has always seemed ironic that this so-called unconventional agent should be pursued experimentally by largely conventional approaches. Although the prior arguments are in some measure persuasive, like Weissmann I find it difficult to accept them without any involvement of nucleic acids.

I do not, on the other hand, find it difficult to envisage an 'unconventional' nucleic acid, one that might deceive investigators into believing in its absence. My suggestion therefore is that we should look for a nucleic acid that doesn't respond chemically or biochemically in a manner expected of it, but genetically still functions as one. The simplest molecules for this would be DNAs containing a phosphotriester — rather than a diester — bond. Such molecules would be chemically stable, but would respond aberrantly to the usual enzymes or chemicals used in their

study, would not migrate as anions in an electrical field, and in any phenol extraction from cellular materials would partition in the organic (not the aqueous) layer, and so on. However, they could still form Watson-Crick base pairs and thus might genetically respond normally. Conceivably, the modification could even be used to anchor the nucleic acid to the prion protein, in the holoprin model. Like Weissmann's model, this suggestion could account "for most if not all experimental observations, does not involve unknown mechanisms and makes testable predictions". But, as he states for his model, "it may also be wrong". To the best of my knowledge, no one has looked for a 'camouflaged' DNA as a factor in spongiform encephalopathies.

BEVERLY E. GRIFFIN

Royal Postgraduate Medical School,
London W12 0NN, UK

Out on a limb

SIR — In 1987, a paper¹ and accompanying News and Views article² were published in *Nature* in which it was suggested that retinoic acid is a morphogen. Since then there have been many reports purporting to confirm this claim. A widely used textbook in cell biology³, for example, states that with regard to determination of the skeletal pattern in the vertebrate limb, "it is very likely that retinoic acid is the natural morphogen".

Previous work had shown that ectopic grafts of an extra zone of polarizing activity (ZPA), a portion of the normal limb bud⁴, as well as resin implants containing retinoic acid⁵, could induce skeletal duplications during limb development. The authors of the 1987 paper¹ showed that ZPA itself is indeed relatively rich in retinoic acid. The claim that retinoic acid is an endogenous limb morphogen could have been tempered, however, by acknowledgement of the long-established fact that the endogenous ZPA is entirely dispensable for normal skeletal development during the period of limb pattern formation⁶. And the implausibility of the notion that a gradient of a single substance, such as retinoic acid, could determine where skeletal elements would form, and their separate identities as well, might also have been considered. When readers of *Nature* learned earlier this year^{7,8} that retinoic acid is almost certainly *not* a morphogen in the developing limb, those outside the field must have been more than a little surprised.

This episode would be of only historical interest if the same phenomenon were not cropping up again with regard to the new 'master regulatory molecules' of the day, the homeobox gene products. Such molecules are clearly of importance in establishing and reinforcing segmental

prepatterns and segment identities in insects. But their roles in pattern formation in nonsegmented organisms such as the nematode *Caenorhabditis elegans* are a matter of speculation⁹, and nothing whatever is known about their causal role in limb skeletal pattern formation. Nonetheless, articles are beginning to appear on the distributions of homeobox gene products in the vertebrate limb (none of which, in fact, actually resembles the skeletal pattern) with claims like "Hox-4 genes probably encode positional information"¹⁰. I believe that it would be preferable to apply the standard that proposed interpretations of developmental pattern formation should attempt to account for why a pattern looks like it does, and not like something else.

STUART A. NEWMAN

Department of Cell Biology and
Anatomy,
New York Medical College,
Valhalla, New York 10595, USA

SLACK REPLIES — How can we know whether a molecule is really acting as a morphogen? Like many others I have spent some time thinking about this and have published a set of possible criteria¹¹. To simplify these somewhat: the substance should have the expected biological activity; it should be present at the right time, place and concentration; and appropriate inhibition should inhibit the effect *in vivo*.

Back in 1987, it seemed to me sufficient that retinoic acid had the right effects and had the predicted distribution *in vivo*. Retinoic acid enters and leaves cells very easily, so a steady-state concentration difference guarantees diffusion from regions of high to regions of low concentration. Since then, other active endogenous retinoids have been discovered¹², it is clear that there are several retinoic acid receptors, and the situation is obviously more complex than it appeared at first sight. A satisfactory inhibition experiment has yet to be devised. So perhaps my News and Views piece² should have been entitled "We think we have a molecule that looks more like a morphogen than any before", but I am not sure that the editors would have liked that as much.

As to Newman's second point, I do not agree. Experimental embryology teaches us that the causal mechanisms in regional specification do *not* 'look like' the final pattern. The model of a morphogen gradient with threshold responses was formulated to account for the embryological data¹³. Now we have some candidate morphogens, the retinoids, and some candidate primary response elements, the Hox-4 genes, and this seems like progress to me. Newman is right to emphasize that we do not understand, even for the primary body

plan of *Drosophila*, the links between the 'serial threshold' arrangement of Hox gene expression and the later visible pattern of cell differentiation. But you can't solve everything at once, and recent progress in developmental biology has been quite spectacular enough.

J. M. W. SLACK

ICRF Developmental Biology Unit,
Department of Zoology,
University of Oxford,
Oxford OX1 3PS, UK

1. Thaller, C. & Eichele, G. *Nature* **327**, 625-628 (1987).
2. Slack, J.M.W. *Nature* **327**, 553-554 (1987).
3. Alberts, B.M. et al. *Molecular Biology of the Cell* 2nd edn p.917 (Garland, New York, 1989).
4. Saunders, J.W. Jr in *Vertebrate Limb and Somite Morphogenesis* (eds Ede, D.A., Hinchliffe, J.R. & Balls, M.) 1-24 (Cambridge Univ. Press, 1977).
5. Tickle, C., Lee, J. & Eichele, G. *Dev Biol.* **109**, 82-95 (1985).
6. Fallon, J.F. & Crosby, G.M. *J. exp. Zool.* **193**, 449-455 (1975).
7. Wanek, N., Gardiner, D.M., Muneoka, K. & Bryant, S.V. *Nature* **350**, 81-83 (1991).
8. Noji, S. et al. *Nature* **350**, 83-86 (1991).
9. Kenyon, C. & Wang, B. *Science* **253**, 516-517 (1991).
10. Izpisua-Belmonte, J.-C. et al. *Nature* **350**, 585-589 (1991).
11. Slack, J. M. W. & Isaacs, H. V. *Development* **105**, 147-154 (1989).
12. Thaller, C. & Eichele, G. *Nature* **345**, 815-819 (1990).
13. Tickle, C., Summerbell, D. & Wolpert, L. *Nature* **254**, 199-202 (1975).

Lead correction

SIR — S. Weiner and G. Goodman (*Nature* **352**, 385; 1991) correctly recommend that, because of inaccuracies, two figures I prepared relating indices of children's lead exposure to their IQ scores should not serve as a basis for policy decisions on acceptable lead levels. They neglected to mention, however, that the inaccuracies have been corrected. The revised versions of these figures also identify the nature and source of the data depicted from each study. I became aware of the need to revise the figures in March, told Weiner this in April and announced it at a meeting on 22-23 April of the US Centers for Disease Control Advisory Committee on childhood lead poisoning. Any reader who wishes copies of the revised figures is invited to contact me.

I apologize to those who may have been misled. But it is important to point out that the basic message conveyed by the figures remains unchanged. The purpose of displaying the data in this manner was to assess the qualitative similarity in the direction and magnitude of the dose-effect relationships reported in various studies, regardless of whether the relationship was significant at $P < 0.05$. The quantitative similarity of the findings from many of the studies depicted has been confirmed by meta-analysis. (H. Needleman and C. Gatsonis *J. Am. med. Ass.* **263**, 673; 1990).

DAVID BELLINGER

Neuroepidemiology Unit,
Gardner House, Children's Hospital,
Boston, Massachusetts 02115, USA

A French *faux pas*

Peter J. Smith

A History of Geology. By Gabriel Gohau. Rutgers University Press: 1991. Pp. 259. \$35 (hbk), \$12.95 (pbk).

WHEN Martin Rudwick, one of the world's most distinguished historians of the Earth sciences, says on the dust jacket of Gohau's book that it is "By far the best short introduction to the history of geology in print — or out of it, for that matter", one must sit up and take notice. But even distinguished historians can be guilty of stating the ineluctable. This is the best short introduction to the history of geology just as the Moon is the best moon we have.

Many books on the history of the Earth sciences have been published over the past 30 years, and some have even covered the whole period of geological studies. Of the latter, however, most have had particular themes (geological time, global models, landforms and so on). A general history of geology was published in 1951 and a disastrous, best-forgotten successor appeared in the early 1980s. In fact, there is only one short up-to-date introduction to the history of geology available and, unfortunately, Gohau's is it.

A brief bibliographical preliminary is necessary here. Gohau's book first appeared in France, in French, in 1987. The present version, beautifully translated by the American historians of geology, Albert V. and Marguerite Carozzi, is a revised and expanded edition of the

the two final chapters on, respectively, continental drift and the birth of the oceans, it's hard to imagine just what American emphasis can possibly have been added, for the bulk of the volume is determinedly chauvinistic.

In terms of topics, there are few surprises. The ideas of the ancients and the luminaries of the Middle Ages, the formation and age of the Earth, neptunism, plutonism, fossils, the beginnings of biostratigraphy, uniformitarianism, catastrophism, mountain-building, the granite problem, the development of the geological timescale, plate tectonics and other, more-or-less expected subjects are there, albeit rather briefly in some cases, as befits a text of only 216 pages. There is also a laudable attempt not merely to recount the facts of history but to trace concepts through time (for example, the cyclic and unidirectional views of Earth history).

But somewhere in the second chapter (in my case) the ghastly truth begins to dawn; with the exception of the ancient Greeks at one end and modern Americans at the other, most of the important discoveries in geology will turn out to have been French. And so it proves. The history of geology is the history of the work of Buridan, Palissy, Descartes, Pallas, de Maillet, Deluc, de Saussure, Buffon, Boulanger, Desmarest, Soula-vie, Deshayes, Cuvier, Brongniart, Lavoisier, Prévost, de Beaumont, Ellenberger and many other Frenchmen, together with the odd foreigner here and there who chipped in the occasional idea to help the French effort along.

William Smith, traditionalists will be horrified to discover, rates four short paragraphs. The first disposes of his life's work in 60 words, the second dismisses him as a mere copier of Cuvier and Brongniart, and the third asserts his lack of reliability — compared to Frenchmen, that is. The fourth is even more interesting. It disingenuously argues that "priority does not really matter in this general history of geology". A twinge of conscience, perhaps?

People such as Hutton and Lyell cannot be rejected quite so easily, although Gohau has a good go at it. Discussion of Hutton is described as a "short detour" and some of his ideas were the "rather naive proposals of a physician-farmer". In any case, "J.-A. Deluc . . . had seen angular unconformities before Hutton" (so much for priority not mattering). As

for Lyell, "To start with, uniformitarianism was not Lyell's invention" but that of Contant Prévost.

White Anglo-Saxon Protestant historians of geology are admittedly on slightly shaky ground here. It is widely acknowledged that early students of the history of geology were overly Anglocentric, having, because of language difficulties, been largely restricted to sources in English (original or translated). Over the past 20 years, however, much, but perhaps not enough, has been

Mary Evans

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Charles Lyell (1797–1875): did he have any real claim to fame?

done to redress this imbalance, with greater credit going to the contributions of French, German, Italian and even American scientists. (Gohau does not even mention G. K. Gilbert, the greatest American geologist of the nineteenth century.)

But the history of the study of the history of geology cannot now be rewritten. To produce, at this stage, an interpretation of the history of geology that cuts right across previous work without explicitly acknowledging revisionism as the chief aim is disgraceful; and to rewrite the history of geology in terms of the work of dozens of obscure (as well as the great) Frenchmen, giving only grudging acknowledgement to a few great geologists of other nations, is merely grotesque.

Perhaps, after all, Gohau is right; perhaps pre-continental-drift geology was indeed an almost entirely French achievement. But if so, the case should first be made in journals subject to peer review, not in an introductory volume aimed largely at the general reader and students. Gohau's book is rather like one of those outrageously unconventional performances of Shakespeare that pop up from time to time; experienced Shakespeare-goers can cope because they have standards by which to judge, but newcomers can be put off for life.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

James Hutton (1726–1797): were his ideas merely "naive proposals"?

original. The updating and the addition of a glossary and references are all to the good, but the claimed "new general emphasis toward the American reader" is all to the mysterious. Lest it be in

Unfortunately, precisely because there is no other recent introduction to the history of geology, newcomers to the field have little choice.

The French Ministry of Culture and Communication gave financial aid towards the translation and publication of *A* (note, not *The*) *History of Geology*. Well, they would, wouldn't they? □

Peter J. Smith is in the Department of Earth Sciences, The Open University, Milton Keynes MK7 6AA, UK.

Close encounters

Clark R. Chapman

Wanderers in Space: Exploration and Discovery in the Solar System. By Kenneth R. Lang and Charles A. Whitney. Cambridge University Press: 1991. Pp. 316. £35, \$54 (hbk); £15.95, \$24.95 (pbk).

THE subject of this book is, literally, everything under the Sun. And a big subject it is too, for we have recently acquired a new understanding of the Solar System thanks to the 'golden age of planetary exploration'. Accounts of the planets published in the early and middle decades of this century seem quaint and mostly wrong today. They were based on the painstaking observations of astronomers, staring at small, blurry images in their telescopes. Astronomers also attached crude instruments (such as spectrometers and bolometers) to their telescopes to analyse planetary radiation.

One can admire the intellectual feat of learning about other worlds from such paltry data. Sometimes the astronomers got it right; often they got it very wrong. But now space-age technology allows cameras and instruments to fly past planets, or even land on their surfaces, and measure them close up. The results of this new technology have been truly revolutionary, changing our concept of our own world and how we came to be here. (The increasing recognition of the role of impacts in shaping planetary history points even to the possible catastrophic end of civilization as we know it.)

The detective skills of astronomers, generalizing from meagre data, have been superseded by the very different skills of geologists, atmospheric scientists, cosmochemists and others who try to make sense of compact-disk libraries full of planetary data. Magellan's radar mapping of Venus is itself testing the limits of how rapidly we can acquire and process vast quantities of data. And



Reaching for Jupiter — a mosaic of Voyager 1 images taken in violet light in 1979. (From *The Journeys of Voyager* by R. Kerrod, published by Prion.)

what we are learning about Venus is being understood not by astronomers but by geoscientists long familiar with discerning essential relationships from the wealth of data provided by our own planet.

With our rapidly changing knowledge of the planets, there is always room for a new summary of what we have learned. And *Wanderers in Space* is a nicely produced, readable and competent attempt at that. There are chapters on each of the planets (and their satellite systems), as well as on comets, asteroids and the origin of the Solar System. Although not explicitly intended as such, this well-organized volume could serve as an introductory text.

With the exception of a series of rather crude, hand-drawn diagrams that purport to summarize each planet, the volume is well illustrated. Paintings and historical illustrations supplement the more usual smattering of graphs and colour photographs taken by planetary spacecraft. Several pages, including four photographs, are devoted to Voyager's 1989 encounter with Neptune.

A problem, though, is that the book is written by two astronomers. I found little that is factually wrong with their accounts of modern planetary science, although sometimes they miss the central points. Often they have intriguing or amusing ways of explaining arcane concepts to the general reader. But we learn more about the authors' education in a field with which they have little professional connection than we learn about planetary science as it is actually researched. Planetary geology and especially cosmochemistry are largely ignored.

The chapter on the Moon is a case in point. Lang and Whitney tell us about what the Moon looks like through a telescope, about its tides, about its gravity field, and about possible theories of its origin. But the geology and geo-

chemistry of the Moon, and the results of the extensive analyses of Moon rocks collected by astronauts, are ignored and trivialized. Indeed, there are several books devoted to lunar basaltic volcanism, but Lang and Whitney reduce the topic to two sentences, one of which bizarrely defines basalt as a kind of rock that on Earth is "associated with lava flows that have solidified in the form of tall columns".

The authors have sometimes failed in the difficult task of keeping fully abreast of rapidly changing fields. It was true a decade ago that only 2,500 asteroids were known but the number is twice as large today. The main developments in planetary science up to the mid-1980s are noted, but are incomplete thereafter. The annotated bibliography is particularly useful until roughly 1989.

In general, the book provides a useful and accessible astronomical view of the Solar System. It should be supplemented by other books that view our world, its environs and its neighbours from the perspectives of the other disciplines that comprise planetary science: space physics, cosmochemistry, planetary geology and geophysics, the atmospheric sciences and even exobiology. □

Clark R. Chapman is in the Planetary Science Institute, 2421 East Sixth Street, Tucson, Arizona 85719, USA.

■ Two related books have recently been published: *Space Commerce* by John L. McLucas (with a foreword by Arthur C. Clarke) relates the history of private enterprise in space and discusses its influence on world space policy and research (Harvard University Press, price £19.95); and *Space Weapons and the Strategic Defense Initiative* in which Crockett L. Grabbe surveys the developments in space weapon systems since SDI, assessing the prospects, weaknesses and potential dangers of strategic defence against nuclear missiles (Iowa State University Press, price \$27.95). □

Underground mysteries

Brian Bertram

The Biology of the Naked Mole-Rat.

Edited by P. W. Sherman, J. Jarvis and A. Alexander. *Princeton University Press*: 1991. Pp. 529. \$65, £40 (hbk); \$24.95, £17.75 (pbk).

LITTLE volcano-shaped mounds on the dry ground of the horn of Africa are the only signs of an astonishing underground community, which has everything to catch the headlines: slavery, incest, bloody fights, nakedness, faeces-eating, temporary castration and remarkable teamwork. A new book at last brings all these goings-on to the surface. Its title inevitably undersells the variety and fascination that it contains within.

The beasts are unattractive at first sight, but in fact they are scarcely ever seen, which is partly why the intricacies of their social system have only started to be unravelled in recent years.

The book, well-edited by the three principal workers on the species, brings together and summarizes nearly two decades of study by some 22 researchers on the species, both in the wild and in perspex tubes in the laboratory. The extent of cooperation among the mole-rat people over the years, forsaking temptations to race one another into print, has been highly laudable and productive, matching the marked cooperation exhibited by their subjects.

All sorts of facts are revealed about these most unusual rodents. Naked mole-rats are pinkly naked with miscellaneous whiskers, almost blind and partly poikilothermic (that is, with a lowerable body temperature). They live totally underground, gnawing their way through very hard soil with their incisors, and working cooperatively to kick the huge amounts of soil backwards and eventually to throw it out onto the surface. They live in colonies of about 70 to 80 individuals, which scuttle along in a tunnel system roughly 3 kilometres in extent, and feed on the large underground food stores of arid country plants. There is relatively little genetic variation between colonies and very little within the clearly inbred colonies.

For small mammals (mean weight 34 grammes), they are extraordinarily long-lived — at least a dozen years. The gestation period is 2.5 months, compared with 2 weeks for a mouse and 2 months for a cat. The young take a year to reach adult size, which depends on many factors, including the presence of a breeding female. She herself is almost twice the average adult size, and particu-

larly elongated.

Naked mole-rats are the most highly social of any species of vertebrate. Only one female in the colony reproduces, and only two or three males. The rest are sterile, their reproductive maturation suppressed by the queen by both chemical and behavioural means. This suppression is lost if the queen is removed or dies, when almost any female has the capacity to compete to replace her.

The queen may retain her reproductive status and monopoly for several years, producing litters of about a dozen young (up to 27) every 2.5 to 3 months. Many other colony members help with their upbringing by carrying, retrieving, grooming, bringing back food and providing faeces.

Social living so cooperative that it involves individuals forgoing their own reproduction is termed eusociality; it is known in the colonial bees and wasps, the termites, in some aphids and now in naked mole-rats. Why it should evolve, in any particular case, and in general, is much debated both in this fascinating book and elsewhere.

It is clear that in their present niche and habitat, naked mole-rats can survive only if living socially. They need companions to share in the huge task of burrowing long distances through a hard substrate to find sparsely distributed food; but, once found, the food can feed many individuals. Group defence against predatory snakes is probably essential. A naked mole-rat out of its own tunnel system is likely to be doomed.

The evolution of sterility is a puzzle — exactly how does natural selection favour the spread of the trait of not breeding? A satisfying partial solution for workers of social bees and wasps was provided by Bill Hamilton in 1964; he pointed out that those species are haplodiploid, so a (female) worker bee would on average be less closely related to her own offspring than she would be to her monogamous mother's. But naked mole-rats are diploid like all other mammals, so the key causes must lie elsewhere.

A queen naked mole-rat, surrounded by overlapping generations of her own long-lived offspring, will be more closely related to her own offspring than she would be to the offspring of her daughters, were they to attempt to breed; it would be in her genetic interest, if breeding by the colony is limited, for her to be the breeder. For her daughters, it does not make so much difference, because a worker mole-rat would share half her genetic material with her own daughter, as she does with her younger sisters produced by her monogamous mother. Thus, in any conflict over who should reproduce, there is an asymmetry in the potential benefit. In addition, the confined tunnel network (virtually a one-

dimensional habitat) makes it difficult for a subordinate to avoid the influence of a dominant attempting to control her.

In contrast with tidy explanations, the species is highly variable in its behaviour. This excellent and eagerly awaited book provides a gold-mine of fact and theory on which to base the discussions it will surely stimulate. □

Brian Bertram is at The Wildfowl and Wetlands Trust, Slimbridge, Gloucester GL2 7BT, UK, and was formerly Curator of Mammals at the Zoological Society of London.

Highway codes

John Armstrong

Intracellular Trafficking of Proteins.

Edited by Clifford J. Steer and John A. Hanover. *Cambridge University Press*: 1991. Pp. 745. £85, \$150.

PROTEINS find their way into all sorts of places inside and outside cells. From a common origin in the cytoplasm, the different polypeptide constituents of each cellular organelle and compartment are sorted and delivered to their destinations. These processes involved in the traffic of protein constitute one of the most exciting areas of current cell biology: progress can be so rapid that even those within the field have trouble keeping abreast of developments. There is therefore a demand for a volume such as this one. How far does the book meet its publisher's claim of being "essential reading for beginning students as well as for established investigators"?

First, any 'beginning student' who can be persuaded to lift this tome, two-handed, off the library shelf will encounter a conventional format, with no pretensions to the attention-grabbing techniques of modern cell-biology textbooks. The overall structure is the familiar multi-authored one, and sadly the usual criticisms of such volumes apply — a variable quality of writing, overlap of material and long production time. The last is perhaps the most worrying in a fast-moving area; indeed, three of the authors point this out. P. Courttoy, in his contribution on the "dissection of endosomes", comments between gritted teeth that since the article was written in 1988, several more reviews of the field have appeared, one by himself (although the chapter is still well worth reading). M. Farquhar in her chapter on the Golgi complex concludes philosophically that the pace of change "will render the useful life of this review remarkably short"; but longer, surely, than the production time of the book?

Conventional presentations of the traffic of proteins begin with the birth of the

protein at the ribosome, but this book starts on the outside, with endocytosis, and stays there for 386 pages. The initial "overview" by M. Willingham and I. Pastan is in fact a restatement of these authors' views: coated vesicles do not exist, the endosome (or "receptosome" in their terminology) forms by swelling and is delivered to the Golgi complex. Diversity of opinion is the lifeblood of science, but how should the beginning student cope with the ensuing chapters in which the coated vesicle's existence is (conventionally) taken for granted? No fewer than 11 chapters are largely devoted to endocytosis, in some cases one ligand at a time, with an inevitably massive amount of redundancy. The tenth chapter is an unusual one on autocrine targeting of hormones, and includes the remarkable proposal that certain growth factors enter the cell by way of endocytic vesicles which "fuse with the nuclear membrane, delivering the hormone to the nucleus". Again, authors have every right to present unorthodox models in print, but nonexpert readers have a greater right to be protected from the ensuing topological havoc.

Eventually the book progresses to the exocytotic pathway. This too has an "overview", which oddly is taken up with a discussion of the assembly of dolichol intermediates in glycoprotein synthesis. The persistent reader is then rewarded with an invigorating piece by V. Lingappa on translocation into the endoplasmic reticulum. The remainder of the secretory pathway is covered in varying depth, and there is an excellent presentation of polarized transport by A. Wandinger-Ness and K. Simons in conclusion. Almost as an afterthought, the rest of the subject of the traffic of proteins — transport from the cytoplasm

— is covered in the last three chapters. Sadly, there is no mention of the chloroplast, an organelle which has transport problems all of its own and which has been the source of one of the most useful ideas in the field. But the contributions on mitochondria (M. Douglas, C. Smagula and W. Chen) and peroxisomes (S. Gould and S. Subramani) are among the best in the book.

Apart from the actual content, several aspects of the book's presentation might have benefited from more attention. For a textbook, several of the diagrams (so important in this subject) are muddled to the point of incomprehensibility; alter-

natively, for a reference book, the index is woefully incomplete. The system of numbered sections breaks down at times, leaving topics apparently adrift on the page, and there are few cross-references between chapters.

The editors of the book have taken on a near-impossible task, but could, in fact, have succeeded had not the healthy babies mentioned above been immersed in some suspiciously murky bathwater. □

John Armstrong is in the Membrane Molecular Biology Laboratory, Imperial Cancer Research Fund, PO Box 123, Lincoln's Inn Fields, London WC2A 3PX, UK.

Principles and processes

Leigh Royden

The Solid Earth: An Introduction to Global Geophysics. By C. M. R. Fowler. Cambridge University Press: 1990. Pp. 472. £40, \$59.50 (hbk); £19.50, \$37.50 (pbk).

HERE at the Massachusetts Institute of Technology, every graduate student of Earth sciences goes through a grisly rite of passage called the 'general examination', which includes a three-hour oral exam. When asked by anxious students for advice on how best to prepare, I usually recommend, in addition to a good night's sleep and a tranquil state of mind, a sound basic textbook that emphasizes the concepts of Earth processes. This fine new geophysics textbook will now be added to the top of my list of recommendations, as it promises to be excellent both for teachers and for those seeking a review of these processes from a geophysical point of view. Indeed, one of the reasons that I like it is that it focuses primarily on the structure and dynamics of the Earth, rather than developing geophysical theory while its application to the Earth remains a secondary concern. As such, it unites many elements that in teaching are all too often ignored because they fall between the traditional disciplines of geophysics and geology. This practice is especially unfortunate because this is an area where many of the recent advances in Earth sciences have been made.

The book begins with a discussion of plate motions on a plane and a sphere, examining triple junctions, propagating ridges and present-day plate motions. A commendable feature is the way in which Fowler intermingles theory and observation in his examples of Earth structures and processes. For the most part this approach succeeds admirably, although a few of the more complex

examples are difficult to follow, particularly in the sections on past plate motions and the continental lithosphere. But in all fairness, many real examples are complex; it may be difficult, if not impossible, to describe in simple terms the evolution of, say, the Indian Ocean.

One of the few places where the book falls short is in the section on the growth of the continents. I would like to have seen this section cast in the light of the Wilson cycle (the repeated rifting and reassembly of the continents), and it would have benefited from more emphasis on the many different and continual reworkings of continental material. The section would then have followed on naturally from the ones on plate motion. Fowler's examples from the Alps and the Himalayas are good, although they lack the succinct theoretical framework that is needed to give them structure. But these are small quibbles. Most of the book is excellent, in both the quality and the clarity of its presentation.

Many of the chapters are reasonably self-contained, making them suitable for a variety of introductory courses. And as a whole, the book will prove to be a superb review text, even for researchers. It will probably not appeal to the casual reader because much of the material requires a certain diligence if it is to be absorbed; a better book in this regard would be *The Earth* by S. Press and R. Siever (W. H. Freeman, 1986). Nor will it supplant *Geodynamics* by D. Turcotte and G. Schubert (Wiley, 1982), which is the cornerstone of introductory geophysical theory and equations, but which fails considerably in the application of these concepts to the Earth. *The Solid Earth* is much needed in that it presents a simple quantitative approach that goes hand-in-hand with an abundant supply of examples illustrating the richness and complexity of the processes that have shaped the Earth. □

Leigh Royden is in the Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

Technical tips

■ In *The Chemical Synthesis of Peptides*, John Hens surveys and details methods for proteins of all kinds, emphasizing the fundamental principles involved. Published by Oxford University Press, price £35.

■ *Antibody Engineering: A Practical Guide*, edited by Carl A. K. Borrebaeck, presents an overview of this rapidly growing field, covering topics from antibody structure to PCR cloning. Published by W. H. Freeman, price £27.95.

■ *Introduction to Flow Cytometry*, edited by James V. Watson, describes the fundamental principles, basic methods and applications of this technique for sorting macromolecules and cells. Published by Cambridge University Press, price is £50, \$79.50. *Quantitative Fluorescence Microscopy* by F. W. D. Rost is a companion book also recently published by Cambridge. Price is £40, \$69.50. *Flow Cytometry* edited by Z. Darzynkiewicz and H. A. Crissman is a comb-bound laboratory manual. Published by Academic, price is £35.50, \$49.95. □

Molecular cloning and characterization of the rat NMDA receptor

Koki Moriyoshi, Masayuki Masu, Takahiro Ishii, Ryuichi Shigemoto*, Noboru Mizuno* & Shigetada Nakanishi

Institute for Immunology and * Department of Morphological Brain Science, Kyoto University Faculty of Medicine, Kyoto 606, Japan

A complementary DNA encoding the rat NMDA receptor has been cloned and characterized. The single protein encoded by the cDNA forms a receptor-channel complex that has electrophysiological and pharmacological properties characteristic of the NMDA receptor. This protein has a significant sequence similarity to the AMPA/kainate receptors and contains four putative transmembrane segments following a large extracellular domain. The NMDA receptor messenger RNA is expressed in neuronal cells throughout the brain regions, particularly in the hippocampus, cerebral cortex and cerebellum.

GLUTAMATE, a major excitatory neurotransmitter in the central nervous system (CNS), has an important role in neuronal plasticity and neurotoxicity¹⁻⁴. The functional diversity of glutamate is reflected by the presence of disparate receptors that can be categorized into two groups, ionotropic and metabotropic¹⁻⁵. Ionotropic receptors contain integral cation-specific ion channels, whereas metabotropic receptors function by intracellular signalling through G proteins¹. Ionotropic receptors are further subdivided into two distinct types: receptors for NMDA (*N*-methyl-D-aspartate), and non-NMDA receptors for α -amino-3-hydroxy-5-methyl-4-isoxazolepropionate (AMPA) and kainate¹. Recently, AMPA/kainate and metabotropic receptors have been characterized by molecular cloning (refs 6-14 and Y. Tanabe *et al.*, manuscript in preparation), and both consist of many different subtypes.

The NMDA receptor has a key role in many functions of glutamate in the CNS. It is essential for inducing long-term potentiation (LTP)^{1,3}, an activity-dependent enhancement of synaptic efficacy that is most pronounced in the hippocampus and which could underlie learning and memory¹⁵. The NMDA receptor also triggers neuronal degeneration and neuronal cell death⁴. Thus, it is thought to mediate many central actions of glutamate, including cognitive processes, memory acquisition, learning and some neurological disorders such as stroke, epilepsy and some neurodegenerative diseases¹⁻⁴. The integral channel of the NMDA receptor allows Ca^{2+} to permeate as well as Na^+ and K^+ , and the increased intracellular Ca^{2+} in neuronal cells is thought to be responsible for evoking both neuronal plasticity and neurotoxicity of glutamate^{1,15}. The NMDA receptor is also thought to have at least five pharmacologically distinct sites¹. These include a glutamate-binding site, a glycine-binding regulatory site, a site in the channel that binds channel blockers such as phencyclidine and related compounds, a voltage-dependent Mg^{2+} -binding site and a Zn^{2+} -binding site. Despite the physiological importance and the well studied characterization of the NMDA receptor, the molecular characteristics of this receptor have remained elusive.

We isolated a functional cDNA clone for the rat NMDA receptor (NMDAR1) from a brain cDNA library using a cloning approach involving a *Xenopus* oocyte-expression system com-

bined with electrophysiology^{13,16}. We report here the structure, electrophysiological and pharmacological properties, and expression pattern of the NMDA receptor.

Cloning and characterization

The NMDA receptor has been extensively characterized electrophysiologically in *Xenopus* oocytes injected with the brain poly(A)⁺ RNA (refs 17-19). We used a cloning strategy combining the oocyte expression system and electrophysiological measurements^{13,16}. Poly(A)⁺ RNA was isolated from rat forebrains and size-fractionated by centrifugation on sucrose gradient. A fraction giving potent electrophysiological response to NMDA (100 μM) was identified in *Xenopus* oocytes measured in a Mg^{2+} -free medium supplemented with glycine (10 μM). A rat forebrain cDNA library was constructed in an RNA expression vector from the response-evoking mRNA fraction. The mixture containing a functional NMDA receptor cDNA clone was identified by testing for the receptor expression in oocytes after they were injected with the mRNAs synthesized *in vitro* from groups of the cDNA mixtures. A single receptor cDNA clone (λN60) was purified by serially subdividing the response-positive cDNA mixture. The λN60 was converted into the plasmid cDNA (pN60) by *in vivo* excision.

TABLE 1 Pharmacology of the cloned NMDA receptor

Compound alone	Response (%)
100 μM NMDA	100
10 μM L-Glutamate	212 $\pm$ 15
100 μM Ibotenate	71 $\pm$ 20
200 μM Quisqualate	44 $\pm$ 10
100 μM L-HCA	85 $\pm$ 19
500 μM Kainate	<5
50 μM AMPA	<5
100 μM tACPD	<5
1 mM GABA	<5
100 μM NMDA (-glycine)	35 $\pm$ 6
Compound plus 100 μM NMDA	
10 μM D-APV	27 $\pm$ 6
1 μM $\pm$ -CPP	47 $\pm$ 7
1 μM CGS 19755	43 $\pm$ 5
50 μM 7-Cl-KYNA	2 $\pm$ 1
1 μM (+)-MK-801	7 $\pm$ 4
100 μM Zn^{2+}	22 $\pm$ 4
100 μM GAMS	97 $\pm$ 2
100 nM JSTX	83 $\pm$ 6
100 μM CNQX	70 $\pm$ 6

Steady-state currents elicited by application of the various agonists and antagonists indicated were recorded from oocytes injected with *in vitro* synthesized pN60 RNA. All compounds were tested in a Mg^{2+} -free standard medium supplemented with 10 μM glycine, except in 100 μM NMDA (-glycine), where glycine was omitted. The magnitudes of responses evoked by various agonists are indicated as percentages of the response elicited by 100 μM NMDA. In the analysis of antagonists, the responses of 100 μM NMDA in the presence of the various antagonists indicated were recorded and are displayed as percentage of the response elicited by 100 μM NMDA. Each value represents the average $\pm$ s.e.m. of at least four recordings in different oocytes. tACPD, *Trans*-1-aminocyclopentane-1,3-dicarboxylate.

The electrophysiological properties of NMDAR1 were first examined in oocytes injected with the mRNA derived from clone pN60. Application of 100 μ M NMDA to mRNA-injected oocytes elicited an inward current of 40–70 nA under voltage clamp at -80 mV in a Mg^{2+} -free medium supplemented with 10 μ M glycine (Fig. 1a). The response showed a rapid initial spike that decayed in a few seconds to a relatively steady level. Neither uninjected oocytes nor water-injected oocytes showed any responses to NMDA. The NMDA response was immediately and almost completely abolished by addition of 100 μ M Mg^{2+} , and the response was rapidly recovered on removal of Mg^{2+} . The omission of glycine also greatly reduced the NMDA response. But this reduction was partial, inconsistent with a previous report indicating the absolute requirement of glycine for the activation of the NMDA receptor in *Xenopus* oocytes injected with rat brain poly(A)⁺ RNA (ref. 18). In our experiments, oocytes injected with brain poly(A)⁺ RNA similarly showed a reduced but significant inward current on removal of glycine. Thus a small amount of glycine could possibly be secreted from residual follicular cells or oocytes in our expression system. An initial spike caused by the addition of NMDA has been reported to involve the activation of a calcium-dependent chloride channel as a result of the NMDA receptor-mediated Ca^{2+} permeation¹⁹. The spike current as well as the steady-state current was enhanced by increasing the concentration of external Ca^{2+} (Fig. 1b). The spike current was also abolished by substitution of Ca^{2+} with Ba^{2+} (Fig. 1b). The results indicate that the cloned receptor can allow Ca^{2+} to permeate.

The voltage dependency of Mg^{2+} inhibition of the NMDAR1 was examined by recording the NMDA response at various holding potentials in the presence and absence of 100 μ M Mg^{2+} (Fig. 1c). The current-voltage curve of the NMDA response

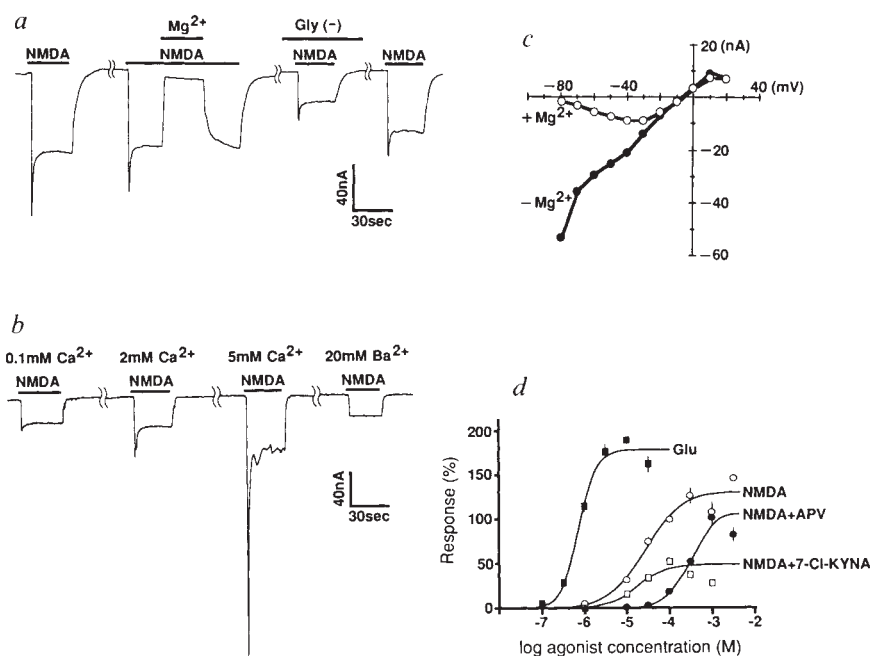
was almost linear between -80 mV and 20 mV in the medium lacking Mg^{2+} . By contrast, the NMDA response was markedly inhibited by the addition of Mg^{2+} at holding potentials between -80 mV and -30 mV. This Mg^{2+} block was released at holding potentials more positive than -20 mV, and the current reversed at about -10 mV regardless of the presence or the absence of Mg^{2+} . The result thus shows that the cloned receptor is blocked by Mg^{2+} in a voltage-dependent manner.

The NMDA receptor can be distinguished from other glutamate receptors by the actions of a number of selective agonists and antagonists. A profile of agonists and antagonists of the NMDAR1 was thus determined electrophysiologically in *Xenopus* oocytes. The agonists L-glutamate, ibotenate, quisqualate and homocysteate (HCA) were effective in inducing electrophysiological responses, L-glutamate being the most potent (Table 1). The values of effective dose for half-maximal response (ED_{50}) of L-glutamate and NMDA were determined to be 0.72 and 27 μ M, respectively (Fig. 1d), in agreement with values obtained in oocytes injected with brain poly(A)⁺ RNA (ref. 20). No (or negligible) responses were observed for other glutamate-related compounds that are specific agonists for different types of glutamate receptors (Table 1).

Various L-antagonists for the NMDA receptor as well as those for non-NMDA receptors were tested for the inhibition of electrophysiological response elicited by the addition of 100 μ M NMDA in *Xenopus* oocytes (Table 1). D-(−)-2-amino-5-phosphonovalerate (D-APV), 3-(±)-2-carboxypiperazin-4-yl-propyl-1-phosphonate (CPP) and CGS 19755, all of which are competitive antagonists for the NMDA receptor², reduced the NMDA response by 50–70%. In the dose-response analysis, 10 μ M D-APV shifted the NMDA response to the right without marked change of the maximal response induced by NMDA application alone, confirming the competitive antagonism of

FIG. 1 Electrophysiological and pharmacological characterization of the NMDAR1 in *Xenopus* oocytes. a, Effects of Mg^{2+} (100 μ M) and glycine on NMDA (100 μ M)-induced current responses were analysed in a *Xenopus* oocyte injected with *in vitro* synthesized pN60 RNA (5 ng) under voltage clamp at -80 mV. Bars show the duration of perfusion of the indicated media; Gly(−), glycine removed from the medium. Downward deflection indicates inward current. b, Effects on NMDA responses (100 μ M) of either increasing concentrations of external Ca^{2+} or replacement of Ca^{2+} with Ba^{2+} were analysed in an mRNA-injected oocyte. c, Voltage dependency of Mg^{2+} block on the NMDA response was analysed by recording electrophysiological responses to 100 μ M NMDA at indicated holding potentials in the presence and absence of 100 μ M $MgCl_2$. d, Dose-response curves of glutamate and NMDA were analysed by measuring steady-state currents after application of varying concentrations of L-glutamate and NMDA. For the analysis of the antagonism of D-APV and 7-Cl-KYNA, the responses to the indicated concentrations of NMDA were measured in the presence of 10 μ M D-APV or 20 μ M 7-Cl-KYNA. The response to application of 100 μ M NMDA alone was taken as 100%. Each point represents the mean $\pm$ s.e.m. of at least four responses recorded in different oocytes.

METHODS. A functional cDNA clone for the NMDA receptor (pN60) was isolated as follows: poly(A)⁺ RNA prepared from the forebrains of male rats (4-week-old) was subjected to centrifugation on sucrose density gradient (5–25%)¹³. A fraction (about 3–5 kb) giving potent electrophysiological response to NMDA in *Xenopus* oocytes was used for the construction of a cDNA library¹³ in λ ZAPII vector. This library was divided into fractions, each containing roughly 3,000 cDNA clones. DNA was extracted from each fraction and subjected to *in vitro* transcription by T7 RNA polymerase in the presence of capping nucleotide¹³. The resultant RNA was injected into oocytes separately. After incubation of oocytes for 3–6 days, they were treated with



2 mg ml^{−1} collagenase for 0.5–1 h and defolliculated manually. Electrophysiological measurements were done in a perfusion medium containing 95 mM NaCl, 2 mM KCl, 2 mM $CaCl_2$, 5 mM HEPES pH 7.5 and 10 μ M glycine under two-microelectrode voltage clamp at -80 mV (ref. 13). A single NMDA receptor cDNA clone (λ N60) was obtained by repeating extraction of DNA, *in vitro* transcription and electrophysiological measurements after stepwise fractionation of the response-evoking cDNA mixture. The plasmid (pN60) was rescued from λ N60 according to the *in vivo* excision procedure. The cDNA insert of pN60 was sequenced in both strands⁴⁴.

ists, non-NMDA receptor antagonists such as γ -D-glutamylaminomethyl sulphonate (GAMS)²¹ and Joro spider toxin (JSTX)²² showed virtually no or only slight inhibition of the NMDA response. 6-Cyano-7-nitroquinoxaline-2,3-dione (CNQX), which acts as a competitive antagonist for non-NMDA receptors³ and for the NMDA receptor noncompetitively²¹, suppressed the NMDA response by about 30%. Thus, the observed electrophysiological and pharmacological properties of the cloned receptor demonstrate that it represents a functional NMDA receptor and support the conclusion that the single protein encoded by the cloned cDNA is sufficient to form an

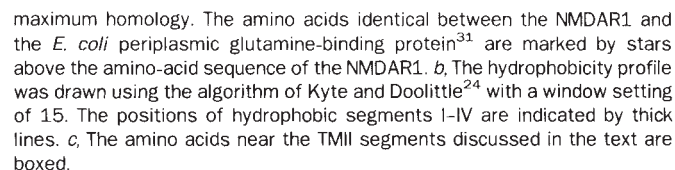


FIG. 3 *a*, Sequence homology comparison between the NMDA receptor (NMDAR) and the AMPA/kainate receptor (GluR1)⁶. (Single-letter amino-acid code.) *b*, A hydrophobicity profile of the NMDAR1. *c*, Amino-acid sequences near the putative TMI segments of the NMDAR1, six subtypes of the rat AMPA/kainate receptors (GluR1-6) and the *Torpedo* nicotonic acetylcholine receptor α subunit (AChR α). *a*, Exact matches and conservative substitutions are shown by bars and colons, respectively. Gaps (—) are inserted for

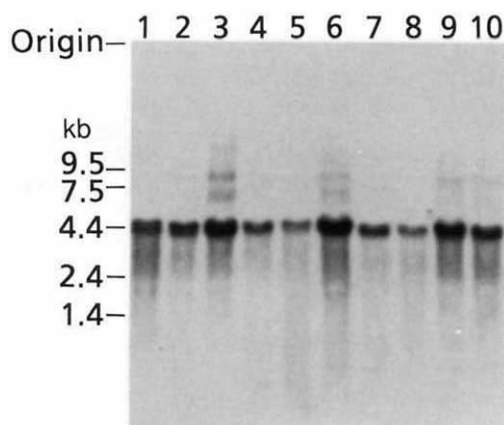


FIG. 4 RNA blot analysis of the rat NMDAR1 mRNA. Total RNAs (10 μ g) analysed¹³ are as follows: lanes 1, whole brain; 2, cerebral cortex; 3, hypothalamus; 4, midbrain; 5, striatum; 6, hippocampus; 7, medulla/pons; 8, spinal cord; 9, olfactory bulb; 10, cerebellum. The cDNA probe and size marker (kb) were the 1,600-bp *Bgl*III fragment of pN60 and the RNA ladder (BRL), respectively.

active receptor/ion channel complex characteristic of the NMDA receptor.

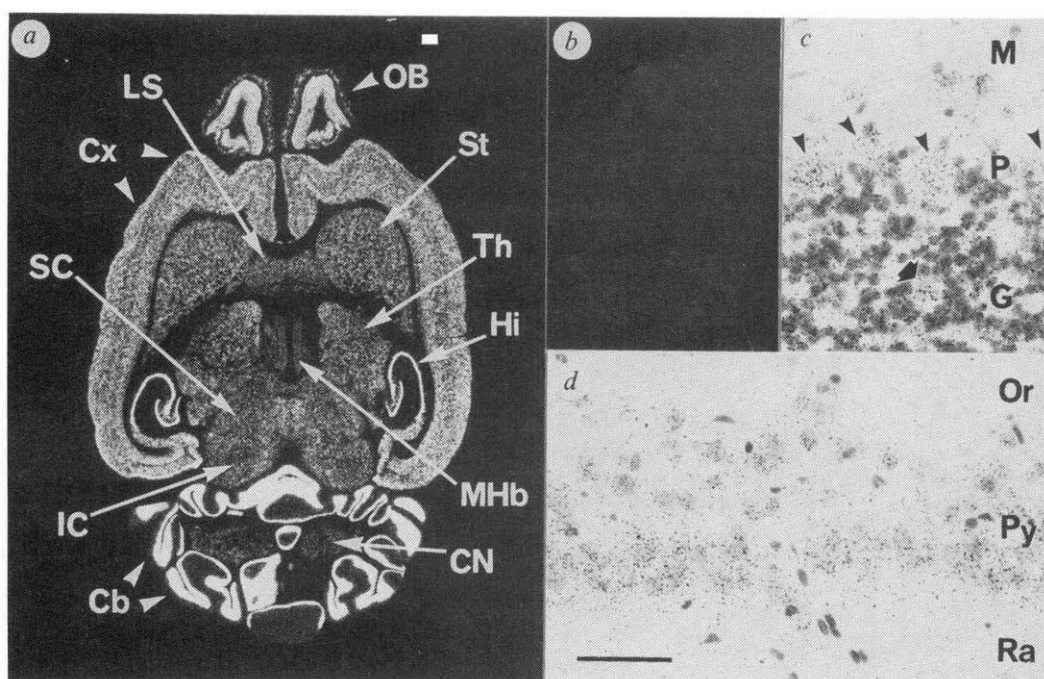
Structure

The nucleotide sequence of the cloned cDNA showed a large open-reading frame of 2,814 base pairs (bp) in the total length of 4,213 bp, the only frame without multiple termination codons (Fig. 2). The nucleotide sequence surrounding the initiation codon agrees well with the consensus sequence²³. The predicted polypeptide consists of 938 amino acids, with a calculated relative molecular mass of 105,500 (M_r 105.5K). The sequence homology analysis revealed a significant similarity between the NMDAR1 and the AMPA/kainate receptors throughout the protein sequences (Fig. 3a). The overall homology between these two receptors is about 22–26%; the subgroups of the AMPA/kainate receptors GluR-4 and GluR5-6, house a similar homology to the NMDAR1 at the corresponding positions. The AMPA/kainate receptors are thought to form a structure in

which a large amino-terminal extracellular domain is followed by four transmembrane segments (TMI-IV)^{6–10} (Fig. 3a). A hydrophobicity analysis²⁴ of the NMDAR1 showed the presence of one hydrophobic sequence at the amino terminus and four hydrophobic segments consisting of more than 20 consecutive uncharged amino-acid residues at the carboxy-terminal half of the protein sequence (Fig. 3b). Thus, taking into consideration the hydrophobicity profile and sequence homology data, we predict a signal peptide at the amino terminus²⁵ and assign four transmembrane domains as depicted in Fig. 2. This assignment of TMI and TMII differs from that originally suggested for the AMPA/kainate receptor⁶ but agrees with the predicted transmembrane domains proposed later for these receptors^{7,9,10}. According to this assignment, the NMDAR1 contains extracellular domains composed of about 540 and 110 amino-acid residues at the amino and carboxy sides of the four transmembrane domains, respectively.

There are several interesting structural features of the

FIG. 5 Localization of the NMDAR1 mRNA in the adult rat brain by *in situ* hybridization. *a*, Negative film image of *in situ* hybridization of a horizontal section. Cb, Cerebellar cortex; CN, deep cerebellar nuclei; Cx, cerebral cortex; Hi, hippocampus; IC, inferior colliculus; LS, lateral septum; MHb, medial habenula; OB, olfactory bulb; SC, superior colliculus; St, striatum; Th, thalamus. *b*, A control hybridization experiment carried out in an adjacent section using the same riboprobe in the presence of 100-fold excess of unlabelled probe. *c*, *d*, Bright-field photomicrographs of emulsion-dipped sections, showing the cerebellar cortex (*c*) and CA1 of the hippocampus (*d*). M, Molecular layer; P, Purkinje cell layer; G, granular layer; arrowheads, Purkinje cells; arrow, a Golgi cell; Or, stratum oriens; Py, stratum pyramidale; Ra, stratum radiatum. Scale bar, 50 μ m.



METHODS. ³⁵S-labelled anti-sense RNA probe covering the 1,431-nucleotide 5'-terminal sequence of pN60 was transcribed and hybridized as described¹³. The sections were exposed to β max film (Amersham) for 3 days (*a*, *b*)

or dipped in NTB2 emulsion (Kodak) diluted 1:1 with distilled water, developed after 2-week exposure and counterstained with cresylviolet (*c*, *d*).

NMDAR1. (1) The TMII segment of ligand-gated ion channels is involved in lining the ion channels²⁷ and negatively charged amino acids flanking the TMII segment are thought to form anion rings that determine ionic selectivity of ligand-gated cation channels^{26,27}. In accordance with the cationic conductance of the NMDA receptor, the putative TMII segment of the cloned receptor is flanked by negatively charged residues (Fig. 3c). In addition, there is a stretch of glutamic acids preceding the TMII segment. These peculiar negative charges located near the inner vestibule of the channel protein could be involved in the voltage-dependent Mg^{2+} block. (2) In the TMII segment of the AMPA/kainate receptors, glutamine, which is substituted for arginine in some receptor subtypes, helps regulate the channel property and Ca^{2+} permeability^{28,29}. The NMDAR1 has asparagine at the corresponding position (Fig. 3c), which may be important for efficient Ca^{2+} permeation. In addition, there are several serine/threonine residues in the TMII segment of the NMDAR1 and, as reported for nicotinic acetylcholine receptors³⁰, some of these may be involved in ion permeation. (3) As pointed out for AMPA/kainate receptors⁹, the amino-terminal extracellular domain (residues 457–543) and the second cytoplasmic loop (residues 750–792) of the NMDAR1 have about 31% and 25% amino-acid identity, respectively, with the periplasmic glutamine-binding protein of *Escherichia coli* (ref. 31; Fig. 3a). The former may participate in binding to glutamate⁹. (4) There are 10 possible *N*-glycosylation sites³² at the extracellular domain (Fig. 2). Moreover, the putative intracellular domains contain consensus phosphorylation sites for Ca^{2+} -calmodulin-dependent protein kinase type II and protein kinase C (refs 33, 34) (Fig. 2). These enzymes might be important in the induction and maintenance of LTP (ref. 35). Our discussion, however, is based mainly on the structural similarity between the NMDAR1 and other ligand-gated ion channels, and additional biochemical evidence is needed for understanding the functional characteristics of the NMDA receptor.

Expression of mRNA

RNA blot analysis showed two adjacent hybridization bands with estimated mRNA sizes of about 4.2 and 4.4 kilobases (kb) (Fig. 4). The RNA blot showed that the NMDAR1 mRNA is widely distributed throughout the brain regions; the highest expression was observed in the hippocampus, hypothalamus and olfactory bulb. *In situ* hybridization revealed that the mRNA is expressed in almost all the neuronal cells throughout the brain regions (Fig. 5a). No particular hybridization signals were observed in parallel control experiments (Fig. 5b). Prominent expression of this mRNA was observed in the cerebellum, hippocampus, cerebral cortex and olfactory bulb. In the cerebellum, high expression was seen in the granular layer (Fig. 5c). In the hippocampus, large amounts of mRNA expression were observed in granule cells and CA4 cells of the dentate gyrus and in pyramidal cells throughout the CA1–CA3 regions (Fig. 5a, d).

Discussion

The single protein encoded by the rat NMDAR1 cDNA clone forms a receptor/ion-channel complex that has electrophysiological and pharmacological properties characteristic of the NMDA receptor: agonist and antagonist selectivity, modulation by glycine, Ca^{2+} permeability, a voltage-dependent channel block by Mg^{2+} , and Zn^{2+} inhibition. The deduced amino-acid sequence of the cloned receptor shows a significant sequence similarity to the members of AMPA/kainate receptors and has four putative transmembrane segments following a large extracellular domain. The NMDA receptor thus belongs to the superfamily of ligand-gated ion channels.

The NMDA receptor/channel complex has been extensively studied by protein purification and other biochemical techniques, but some contradictory results have been reported. The complex purified by affinity chromatography has been reported

to contain four subunits with M_s of about 67, 57, 46 and 33K^{36,37}, which seem to be required for the activity of the NMDA-induced cation influx in reconstituted liposomes³⁷. The sizes of these subunits are, however, much smaller than those determined by radiation inactivation analysis³⁸; the molecular target sizes of bindings of L-glutamate, *N*-(1-(2-thienyl)cyclohexyl)-3,4-piperidine (TCP, an NMDA-receptor channel blocker) and glycine were reported to be about 120K. Furthermore, a polypeptide with an M_r of about 120K has been identified by photoaffinity labelling of the phencyclidine site of the NMDA receptor complex³⁹. These latter sizes agree with that of our cloned receptor. In the family of AMPA/kainate receptors, there is an additional and different type of the receptor which is much smaller than the other AMPA/kainate receptors^{40,41}. This protein, called the kainate-binding protein, has four transmembrane domains and shows a significant sequence homology with the AMPA/kainate receptors. But no functional channel activity was detected by the expression of the cDNA for the kainate-binding protein in a heterologous expression system⁴⁰, and the function of this binding protein is unknown. It is unclear whether the purified small subunits of the NMDA receptor complex represent different types of this receptor family or contribute as regulatory subunits of the NMDA receptor.

In situ hybridization indicated that the NMDAR1 mRNA is distributed in almost all neuronal cells throughout the brain regions. Overall, the spatial pattern and extent of expression of this mRNA are consistent with autoradiographic studies using several radiolabelled ligands of the NMDA receptor⁴². But there are some disparities between the ligand binding sites and the mRNA expression sites. For example, there is less ligand binding in the hippocampal CA3 region than in the CA1 region⁴³, but the NMDAR1 mRNA is equally abundant in both regions. The brain-stem region is low in ligand binding⁴³ but mRNA expression is considerable. This discrepancy may result from different cellular localizations of the expressed receptor and its mRNA in neuronal cells, or from the possible multiplicity of NMDA receptor subtypes.

The existence of multiple subtypes of AMPA/kainate receptors has been confirmed by molecular cloning^{7–12}. Notably, the AMPA/kainate receptor subtypes form heteromeric oligomers which are different in electrophysiological properties and ionic selectivity^{8–10,29}. The expression patterns of the multiple receptors are also different^{7,8,11,12}, and the formation of heteromeric oligomers of different subtypes of the receptor seems to provide the electrophysiological and pharmacological heterogeneity of the AMPA/kainate receptors observed in neuronal cells^{8–10,29}. As in *Xenopus* oocytes the depolarization induced by the RNA synthesized *in vitro* from our cDNA clone was much smaller than the one observed with forebrain poly(A)⁺ RNA, it is likely that additional subtypes of the NMDA receptor may exist. The cloned NMDA receptor can thus lead to a better understanding of the molecular nature of the NMDA receptors and will also provide clues towards clarifying the mechanisms underlying the diverse functions of the NMDA receptors in normal and pathological conditions. □

Received 10 October; accepted 22 October 1991.

- Monaghan, D. T., Bridges, R. J. & Cotman, C. W. *A. Rev. Pharmac. Tox.* **29**, 365–402 (1989).
- Watkins, J. C., Krosgaard-Larsen, P. & Honore, T. *Trends pharmacol. Sci.* **11**, 25–33 (1990).
- Collingridge, G. L. & Singer, W. *Trends pharmacol. Sci.* **11**, 290–296 (1990).
- Meldrum, B. & Garthwaite, J. *Trends pharmacol. Sci.* **11**, 379–387 (1990).
- Schoepp, D., Bockaert, J. & Sladeczek, F. *Trends pharmacol. Sci.* **11**, 508–515 (1990).
- Hollmann, M., O'Shea-Greenfield, A., Rogers, S. W. & Heinemann, S. *Nature* **342**, 643–648 (1989).
- Keinänen, K. *et al. Science* **249**, 556–560 (1990).
- Boulter, J. *et al. Science* **249**, 1033–1037 (1990).
- Nakanishi, N., Shneider, N. A. & Axel, R. *Neuron* **5**, 569–581 (1990).
- Sakimura, K. *et al. FEBS Lett.* **272**, 73–80 (1990).
- Werner, P., Voigt, M., Keinänen, K., Wisden, W. & Seeburg, P. H. *Nature* **351**, 742–744 (1991).
- Egebjerg, J., Bettler, B., Hermans-Borgmeyer, I. & Heinemann, S. *Nature* **351**, 745–748 (1991).
- Masu, M., Tanabe, Y., Tsuchida, K., Shigemoto, R. & Nakanishi, S. *Nature* **349**, 760–765 (1991).
- Houamed, K. M. *et al. Science* **252**, 1318–1321 (1991).
- Madison, D. V., Malenka, R. C. & Nicoll, R. A. *Rev. Neurosci.* **14**, 379–397 (1991).
- Masu, Y. *et al. Nature* **329**, 836–838 (1987).
- Verdoorn, T. A., Kleckner, N. W. & Dingledine, R. *Science* **238**, 1114–1116 (1987).
- Kleckner, N. W. & Dingledine, R. *Science* **241**, 835–837 (1988).

19. Leonard, J. P. & Kelso, S. R. *Neuron* **4**, 53–60 (1990).
20. Verdoorn, T. A. & Dingleline, R. *Molec. Pharmacol.* **34**, 298–307 (1988).
21. Lodge, D. & Johnson, K. M. *Trends pharmacol. Sci.* **11**, 81–86 (1990).
22. Saito, M. *et al. Brain Res.* **481**, 16–24 (1989).
23. Kozak, M. *Nucleic Acids Res.* **15**, 8125–8148 (1987).
24. Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105–132 (1982).
25. von Heijne, G. *Eur. J. Biochem.* **133**, 17–21 (1983).
26. Imoto, K. *et al. Nature* **335**, 645–648 (1988).
27. Betz, H. *Neuron* **5**, 383–392 (1990).
28. Verdoorn, T. A., Burnashev, N., Monyer, H., Seeburg, P. H. & Sakmann, B. *Science* **252**, 1715–1718 (1991).
29. Hume, R. I., Dingleline, R. & Heinemann, S. F. *Science* **253**, 1028–1031 (1991).
30. Changeux, J.-P., Giraudat, J. & Dennis, M. *Trends pharmacol. Sci.* **8**, 459–465 (1987).
31. Nohno, T., Saito, T. & Hong, J. *Molec. Gen. Genet.* **205**, 260–269 (1986).
32. Hubbard, S. C. & Ivatt, R. J. A. *Rev. Biochem.* **50**, 555–583 (1981).
33. Kemp, B. E. & Pearson, R. B. *Trends biochem. Sci.* **15**, 342–346 (1990).
34. Kishimoto, A. *et al. J. biol. Chem.* **260**, 12492–12499 (1985).
35. Malinow, R., Madison, D. V. & Tsien, R. W. *Nature* **335**, 820–824 (1988).
36. Ikin, A. F., Kloog, Y. & Sokolovsky, M. *Biochemistry* **29**, 2290–2295 (1990).
37. Ly, A. M. & Michaelis, E. K. *Biochemistry* **30**, 4307–4316 (1991).
38. Honoré, T. *et al. Eur. J. Pharmacol.* **172**, 239–247 (1989).
39. Sonders, M. S. *et al. J. biol. Chem.* **265**, 6776–6781 (1990).
40. Wada, K. *et al. Nature* **342**, 684–689 (1989).
41. Gregor, P., Mano, I., Maoz, I., McKeown, M. & Teichberg, V. I. *Nature* **342**, 689–692 (1989).
42. Young, A. B. & Fagg, G. E. *Trends pharmacol. Sci.* **11**, 126–133 (1990).
43. Monaghan, D. T. & Cotman, C. W. *J. Neurosci.* **5**, 2909–2919 (1985).
44. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).

ACKNOWLEDGEMENTS. We thank A. Uesugi for photographic assistance. This work was supported in part by the Ministry of Education, Science and Culture of Japan and the Yamanouchi Foundation for Research on Metabolic Disorders.

Structure of Sindbis virus core protein reveals a chymotrypsin-like serine proteinase and the organization of the virion

Hok-Kin Choi*, Liang Tong*, Wladek Minor*, Philippe Dumas*†, Ulrike Boege‡
Michael G. Rossmann*§ & Gerd Wengler¶

* Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907, USA

‡ Department of Biochemistry, University of Alberta, Edmonton, Alberta T6G 2H7, Canada

¶ Institut für Virologie, Justus-Liebig-Universität Giessen, Frankfurter Strasse 107, D-6300 Giessen, Germany

Sindbis virus consists of a nucleocapsid core surrounded by a lipid membrane through which penetrate 80 glycoprotein trimers. The structure of the core protein comprising the coat surrounding the genomic RNA has been determined. The polypeptide fold from residue 114 to residue 264 is homologous to that of chymotrypsin-like serine proteinases with catalytic residues His 141, Asp 163 and Ser 215 of the core protein positioned as in other serine proteinases. The C-terminal tryptophan remains in the P₁ substrate site subsequent to the autocatalytic *cis* cleavage of the capsid protein, thus rendering the proteinase inactive. Model building of the Sindbis core protein dimer shows that the nucleocapsid is likely to have *T* = 4 quasisymmetry.

TOGAVIRUSES are small enveloped vertebrate viruses containing a single-stranded RNA genome of positive sense. Most members of this family replicate in and are transmitted by arthropods, causing various serious, frequently fatal, diseases such as encephalitis, fever, arthritis and rash¹. Alphaviruses are mosquito-borne togaviruses of which Sindbis virus is the prototype. The virus particles have a relative molecular mass (*M_r*) of about 40 × 10⁶ (refs 2, 3). Their genomic RNA (*M_r*, 4.2 × 10⁶) is surrounded by a protein capsid of either 180 (refs 3–6) or 240 (refs 4, 7, 8) subunits of identical amino-acid sequence arranged with icosahedral symmetry. The Sindbis virus nucleocapsid cores have an external radius of about 205 Å^{3,9} and are surrounded by a lipid bilayer, roughly 40 Å thick¹⁰ (Fig.

1*a*). The membrane is penetrated by 80 glycoprotein 'spikes' which are about 75 Å long and are arranged^{11–13} in a *T* = 4 (Fig. 1*b*) surface lattice¹⁴. Each spike consists of three copies of the E₁ (439 amino acids) and E₂ (423 amino acids) glycoproteins. In Semliki Forest virus, unlike Sindbis virus, three copies of the non-membrane-spanning glycoprotein E₃ (64 amino acids) participate in the formation of the external head of the spike. The spike contains the recognition site for the host cell receptor¹⁵.

The complete sequence of Sindbis 49S RNA has been determined¹⁶ and correlated with the chemically derived amino-acid sequence of the nucleocapsid Sindbis core protein (SCP)¹⁷. Sindbis RNA encodes two polypeptides, a non-structural p270 polypeptide translated directly from the genomic RNA and a structural p130 polypeptide translated from the subgenomic messenger RNA. Both are post-translationally cleaved into their component proteins by viral and cellular proteinases¹⁸. The order of proteins in the p130 polypeptide is SCP-E₃-E₂-E₁. The core protein is released as the first cleavage product, probably as a result of autocatalysis by the core protein itself^{19–21}, and then has no further proteolytic activity¹⁵. Processing and glycosylation of the spike proteins is slower and occurs primarily in the Golgi apparatus. The assembled spikes are united with the cores on the inside of the cellular membrane during budding and release of the mature virions^{22,23}.

Boege *et al.*²⁴ pointed out that the Gly-Asp-Ser-Gly sequence (residues 213–216) in the capsid protein corresponds to the highly conserved sequence around the catalytically essential Ser 195 in chymotrypsin-like serine proteinases. (Residues in italic numbers are for SCP, whereas other residue numbering uses the chymotrypsin-like proteinase nomenclature.) Hahn *et al.*²⁵ characterized three temperature-sensitive mutants that inhibited capsid assembly at the nonpermissive temperature. They suggested that His 141, Asp 147 and Ser 215 constituted the catalytic triad similar to that required in other serine proteinases. Melancon and Garoff²⁶ used site-specific mutagenesis to verify that the residue homologous to Ser 215 in Semliki Forest virus was essential for catalysis. Hahn and Strauss²⁷ then used site-specific mutagenesis of an infectious clone and identified His 141, Ser 215 and probably Asp 163, rather than Asp 147, as

† Present address: Laboratoire de Cristallographie Biologique, Institut de Biologie Moléculaire et Cellulaire, Centre National de la Recherche Scientifique, 15 rue René Descartes, 67084 Strasbourg Cedex, France.

§ To whom correspondence should be addressed.

the catalytic triad essential for the autocatalytic processing of the capsid protein. They also found that substitution of Ser 215 by cysteine or threonine, of His 141 by arginine and of Asp 163 by histidine did not inhibit processing of the polyprotein, but did inhibit assembly of infectious particles.

The sequence homology with serine proteinases suggests a common evolutionary origin of the capsid protein and chymotrypsin-like serine proteinases²⁵. On the other hand, Fuller

and Argos⁶ suggested sequence similarities between the capsid protein and other viral capsid proteins, in particular VP3 of picornaviruses. Their suggestion was supported by their (probably erroneous) demonstration of $T=3$ capsid symmetry. As many plant and animal viruses have $T=3$ or pseudo $T=3$ symmetry and similar eight-stranded antiparallel β -barrel folds in their capsid proteins, it was reasonable to suggest, particularly when reinforced by amino-acid sequence similarity, that the tertiary structure of SCP should be similar. Such a structure also placed in juxtaposition the amino acids of the catalytic triad, of which Asp 147 is now known to have been misidentified.

Structure determination

Three different types of SCP crystals, which diffract fairly well, have been reported previously²⁸. In each case, the Matthews coefficient, V_M (ref. 29), was only about $1.6 \text{ \AA}^3$ per dalton, assuming M_r of $\sim 29,000$ per SCP monomer. However, if the very basic, disordered, initial 113 residues are omitted from the molecular weight calculation, V_M has the far more normal value of $2.8 \text{ \AA}^3$ per dalton. We report here the structure determination of the ordered component of SCP, corresponding to residues 114 to 264, by using simultaneously X-ray diffraction data of the tetragonal type-2 crystals ($a = 57.0$ and $c = 109.8 \text{ \AA}$ in space-group $P4_32_12$) and the monoclinic type-3 crystals ($a = 38.8$, $b = 79.7$, $c = 60.8 \text{ \AA}$, $\beta = 102.2^\circ$ in space-group $P2_1$). The autocatalytic properties and the organization of SCP into nucleocapsids are then examined on the basis of the SCP structure.

SCP was extracted from nucleocapsids by disrupting the cores with acetic acid and urea²⁴. SCP crystals were then grown as described by Boege *et al.*²⁸. The Matthews coefficient, V_M , indicated that there can be only one SCP monomer per asymmetric unit in the type-2 crystals and that there are probably two monomers per asymmetric unit in the type-3 crystals. Native and heavy-atom derivative X-ray diffraction data were collected on a Siemens area detector and processed with the XDS program³⁰. The heavy-atom sites were solved with difference Pattersons and difference electron-density maps. The heavy atoms in type-3 crystals could be segregated into two groups, such that each group had very similar interatomic distances. These same distances were maintained in the equivalent heavy-atom compounds prepared for type-2 crystals. The rotational relationships between the two groups in type-3 crystals and one group in type-2 crystals were consistent with self- and cross-rotation functions³¹. It was then clear that there was a non-crystallographic dimer axis in type-3 crystals and that this axis corresponded to the crystallographic 2-fold axis along $[110]$ in type-2 crystals.

The heavy-atom derivative data were used to compute multiple isomorphous replacement phases to $3 \text{ \AA}$ resolution for both type 2 and 3 crystal forms. Comparison of the two densities, after establishing the correct hand of the molecule, showed that the space group of type-2 crystals was $P4_32_12$, rather than $P4_12_12$. By averaging the electron density between the two crystal forms, it was possible to determine an envelope for the SCP dimer. Cycles of electron-density calculations and averaging of density between crystals⁶⁴ produced an interpretable electron density map at $3.0 \text{ \AA}$ resolution (L.T., H.K.C. and M.G.R., unpublished results). The electron density is consistent with the known amino-acid sequence of SCP and the chemical nature of the heavy-atom substitutions. Crystallographic refinement of both structures is in progress and the current R -factor is 22% for type-2 crystal data between 6 and $3 \text{ \AA}$ resolution. Preliminary, unrefined coordinates are deposited with the Brookhaven Protein Data Bank.

The structure

The SCP dimer is $\sim 75 \text{ \AA}$ long, but only $\sim 20 \text{ \AA}$ wide and $\sim 34 \text{ \AA}$ thick. The monomer is very roughly spherical, but the dimer is

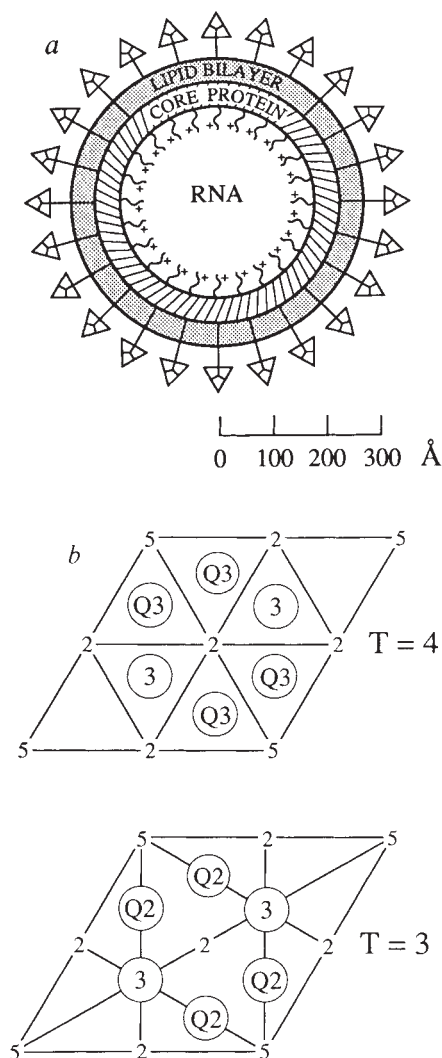


FIG. 1. a, Diagrammatic cross-section of the Sindbis virus structure. The virion has 80 trimeric glycoprotein spikes arranged in a $T=4$ surface lattice. The spikes penetrate a lipid bilayer and are anchored in the core protein capsid. The basic N-terminal segments of the Sindbis core protein (SCP) are associated with the Sindbis virus genomic RNA. b, Organization of the trimeric glycoprotein spikes (circles) in $T=4$ and $T=3$ surface lattices. Numerals indicate icosahedral or quasi-icosahedral axes. The latter are labeled with a 'Q'. The diagram covers 2 of the 20 icosahedral faces centred on 3-fold axes and bounded by 5-fold vertices. Note that some of the trimeric spikes would be associated with quasi-2-fold axes for a $T=3$ lattice. Some of the quasi-2-fold and quasi-3-fold axes have been omitted in order to emphasize the organization of spikes with respect to the core. An icosahedron is constructed out of 60 identical structures which each have identical environments. Each of the 60 structures can consist of one or more proteins. Larger viruses either code for more than one capsid protein or relax the conditions of exact equivalence between covalently identical proteins. Nevertheless, if a quasi-identical environment is to be maintained, then the possible arrangements are restricted and can be expressed by the triangulation number T (ref. 14). In a $T=3$ virus there are 180 covalently identical proteins, whereas in a $T=4$ virus 240 such subunits make up the viral capsid. These arrangements differ in the organization of quasi-2-, 3- and 6-fold local symmetry operators on the capsid surface (Fig. 1b).

long with the monomer-monomer contacts restricted to a narrow neck made up of residues 185 to 190 in a β -strand as well as residue 222 (Fig. 2a). Each SCP subunit consists of 264 amino acids of which the first 113 are far more variable than the remaining 151 residues among alphaviruses and contain an exceptionally large number of basic residues as well as prolines. Furthermore, there are no aromatic amino acids between residues 22 and 115. Plant virus capsid proteins also frequently contain N-terminal basic regions, similar to that in SCP, which are intimately associated with RNA³²⁻³⁴. The SCP polypeptide chain could be followed without difficulty from Arg 114 to the C-terminal Trp 264 (Fig. 2b). The first 113 residues, as expected, are mostly disordered and account for much of the 'solvent' density. However, there is some electron density, corresponding to about 7 residues, whose assignment remains unknown.

There is a striking similarity of the monomer fold in SCP to mammalian serine proteinases of the chymotrypsin family³⁵ or bacterial serine proteinases such as α -lytic proteinase³⁶ (Fig. 3). The polypeptide is folded into two similar 'Greek key' β -barrel

domains with the substrate site situated between the domains. The ordered part of SCP starts at the beginning of the first β -strand of the first domain (βA_1) (see Fig. 2b for a nomenclature of the secondary structural elements). In comparison with chymotrypsin and α -lytic proteinase, the βA_1 - βB_1 corner is shorter in SCP. Chymotrypsin residue 54 in βC_1 is invariably a serine or threonine hydrogen-bonded to the carbonyl of Gly 43. In SCP, however, this residue is Val 136. The essential His 141 is situated in an α -helix (αA) as it is in chymotrypsin. Most of the 'calcium-binding loop' (residues 65-84 in chymotrypsin) is missing in SCP as well as in α -lytic proteinase. Similarly, the loop between βE_1 and βF_1 is missing in both SCP and α -lytic proteinase. In general, the connection between the first and second domain is variable. In SCP it is shorter and different from that in chymotrypsin and α -lytic proteinase.

In the second domain, the connection between βA_2 and βB_2 , the autolysis loop, is much shorter in both SCP and α -lytic proteinase compared with chymotrypsin (residues 145-154). The 'methionine loop' (residues 164-182) is completely missing in



FIG. 2 The SCP structure. *a*, Stereoscopic view of the C_α polypeptide trace for the SCP dimer. *b*, Ribbon drawing of SCP monomer with secondary structure nomenclature (compare alignments in Fig. 5). Amino-acid numbers are shown at strategic positions and for the catalytic triad. (Drawn by a program package described by Kraulis⁶¹.)

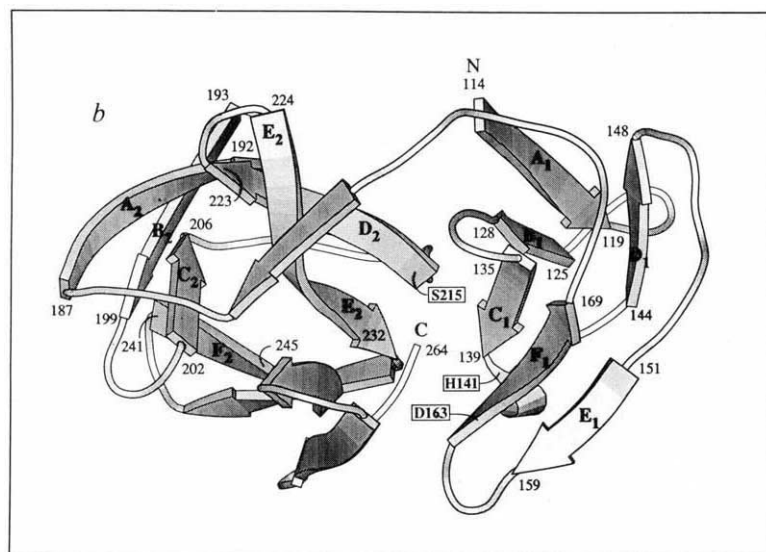
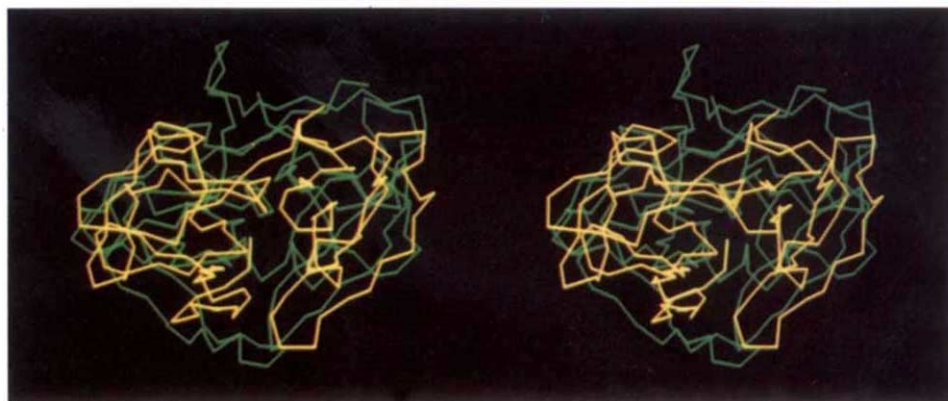


FIG. 3 Superposition of the C_α tracing of SCP (yellow) onto α -lytic proteinase (green). The greater similarity of SCP and α -lytic proteinase to each other than to chymotrypsin is the result of a similar pattern of deletions with respect to the larger chymotrypsin molecule.



SCP but exists in both chymotrypsin and α -lytic proteinase. The βC_2 to βD_2 corner is similar in SCP and chymotrypsin compared with α -lytic proteinase. This section contains the highly conserved sequence Gly-Asp-Ser-Gly (residues 213–216) recognized by Boege *et al.*²⁴ The βD_2 to βE_2 loop is similar and shorter in SCP and α -lytic proteinase compared to chymotrypsin. The βE_2 – βF_2 corner is shorter by one and two residues compared to chymotrypsin and α -lytic proteinase, respectively, and follows a different path. Instead of the final α -helix C (consisting of 16 residues) found in other chymotrypsin-like serine proteinases, the SCP polypeptide follows a different path with its C terminus in the active site.

There are no disulphide bonds in SCP as in most other cytoplasmic proteins. In contrast, α -lytic proteinase has two S–S bonds (one in common with chymotrypsin) and chymotrypsin has four, with a further S–S bond in its zymogen. In most parts, the structure of SCP resembles more closely that of α -lytic proteinase (Fig. 3). Both these proteins have shorter turns and are more compact.

SCP as a proteinase

The structure of SCP confirmed the identification²⁷ of Ser 215, His 141 and Asp 163 as the essential catalytic triad, which has

the same three-dimensional relationship as in other serine proteinases. The C terminus of SCP lies in the substrate-binding site with the terminal Trp 264 in the P_1 position (Fig. 4a and b) and is, therefore, consistent with the inhibition of SCP after the *cis* cleavage of the polypeptide. The protein is an 'enzyme' for only one reaction, after which it exhibits the various functions required of a viral capsid protein.

The active site of SCP has been superimposed on α -lytic proteinase to which an Ala-Pro-Val derivatized peptide inhibitor had been bound³⁷. This showed that the C terminus of SCP binds to itself at the same site and in the same peptide direction as do substrates to other serine proteinases (Fig. 4b), including the differently folded subtilisin molecule. Superposition of the SCP active centre residues His 141, Asp 163, Gly 213, Asp 213, Ser 215, Gly 216, Leu 231 onto the equivalent α -lytic proteinase residues His 57, Asp 102, Gly 193, Asp 194, Ser 195, Gly 196, Ser 214 gave a root-mean-square (r.m.s.) deviation for C_α atoms of 0.8 Å. The α -lytic proteinase inhibitor C_α atoms have an r.m.s. displacement of 1.5 Å with respect to the final three residues of SCP. Superposition of all topologically equivalent residues³⁸ onto α -lytic proteinase and onto α -chymotrypsin gave r.m.s. deviations of 3.02 and 2.56 Å for 100 and 109 superimposed C_α atoms, respectively (Fig. 5).

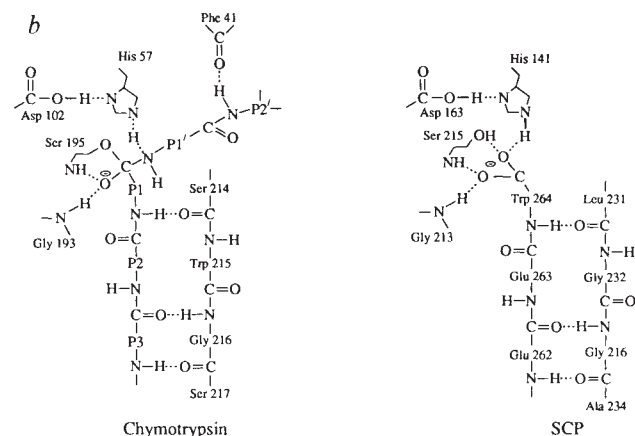
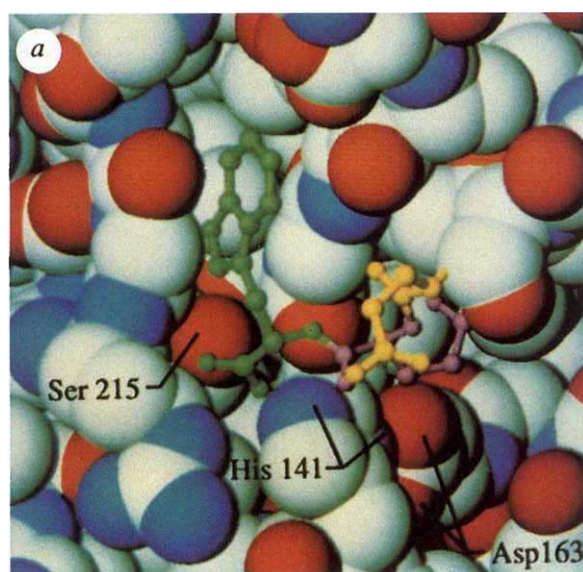


FIG. 4 a, The C-terminus Glu-Glu-Trp (ball and stick model) in the active centre of SCP (space-filling model). Labelled are residues His 141, Asp 163 and Ser 215. (Drawn by a program package described by Connolly⁶².) b,

Diagrammatic representation of the active centres of chymotrypsin and SCP. (Based on a drawing by Liljas and Rossmann⁶³.)

Residues that are entirely conserved between chymotrypsin, α -lytic proteinase and SCP (Fig. 5) include the catalytic triad (His 141, Asp 163 and Ser 215), as well as the sequence Gly-Asp-Ser-Gly around the essential serine. The main-chain amino group of Gly 213 (193 of chymotrypsin) forms an oxyanion hole for the substrate carbonyl preceding the scissile bond. Asp 194 in chymotrypsin, corresponding to Asp 214 in SCP, forms an ion pair with the N terminus of the chymotrypsin polypeptide chain when the mammalian zymogen is activated by cleavage at residue 16. This rearranges the specificity pocket and permits binding of the substrate³⁹. The corresponding ion pair in the bacterial enzymes occurs between Asp 194 and Arg 138. It may, therefore, be significant that in SCP Asp 214 is associated with His 193 and Arg 217. SCP, however, is active immediately after synthesis at the ribosome and does not require any activation step. The highly conserved glycines at 193 and 196 in chymotrypsin may be necessary for steric reasons to build the appropriate conformation around the essential serine residue. Serine 214 of chymotrypsin (Fig. 5), the fourth member of the 'catalytic tetrad', is completely conserved in all other known chymotrypsin-like proteinases. This residue is in the active centre, hydrogen bonds to the catalytic Asp 102 of chymotrypsin and is part of the polypeptide which forms an antiparallel sheet with the substrate (Fig. 4b). The corresponding residue is, however, a leucine in SCP. Other highly conserved residues in SCP are Leu 169, Gly 187 and Gly 233 (Fig. 5), possibly for the purpose of folding.

In trypsin the major specificity pocket at P₁ has Asp 189 at its end to create recognition for arginine or lysine. This residue is a serine in chymotrypsin and a valine in SCP. Other residues surrounding this specificity pocket in SCP are Gly 210, Gly 211, Gly 213, Gly 232 and Gly 233. Thus, the general effect is to create a fairly hydrophobic pocket with ample space to recognize the C-terminal Trp 264. The residues preceding Trp 264 (Glu 262 and Glu 263, corresponding to sites S₃ and S₂) point into solution and are therefore not specifically recognized. In other chymotrypsin-like serine proteinases, P₂ is maintained by conserved

aromatic residues 94 and 171. The corresponding loops containing these residues are missing in SCP, however, and thus there is no recognition of residues P₃ and P₂.

The change in direction of the C terminus of SCP compared with chymotrypsin and α -lytic proteinase leaves the essential Asp 163 exposed to solvent. In other serine proteinases this aspartic acid is unusual in that it is buried, possibly to help polarize the essential histidine for donation of a proton to the substrate during catalysis. It is surprising, therefore, that the essential aspartic acid is exposed in SCP, which suggests that it might be buried when cleavage of the polypeptide occurs while still attached to the ribosome⁴⁰.

Proposed structure of Sindbis cores

Type-2 and type-3 crystals are formed from essentially the same dimer. Similarity of cell dimensions and space-group symmetry²⁸ between type-1 and type-3 crystals suggests that the same dimer also occurs in type 1. Coombs and Brown⁸ show that monomers in the cores can be crosslinked to form dimers with bifunctional agents of the appropriate length. Thus, it is reasonable to assume that the dimer found in the current crystallographic analysis is essential for the assembly of cores. If one attempts to build these dimers into a $T=3$ assembly, there would be significant density at the 2-fold and quasi-2-fold axes, inconsistent with the electron-microscopy results of Fuller³. Furthermore, attempts at building a $T=3$ capsid, using the dimer axes as icosahedral and quasi-icosahedral 2-fold axes, show that the dimers are not long enough to make contact with one another when positioned so as to make particles with an external radius of 205 Å. On the other hand, if the dimer is used to construct a $T=4$ surface lattice (Fig. 6a), a reasonable packing arrangement results which is also consistent with the electron-microscopy reconstruction (Fig. 6b). There are two types of spikes, which are covalently identical but conformationally different^{41,42}. Kääriäinen and Söderlund¹² have shown that 20 of these trimeric spikes are associated with the icosahedral 3-fold axes of the nucleocapsid core, whereas the other 60 would be



FIG. 5 Alignment of chymotrypsin (CHYT) and α -lytic proteinase (α LP) amino acid sequences³⁶ with Sindbis core protein (SCP) on the basis of structural comparisons. Also shown is the alignment of polio 3C proteinase with chymotrypsin based on the results of Gorbalenya *et al.*⁵⁹. Boxed residues

are entirely conserved between CHYT, α LP and SCP. The catalytic triad residues are marked with an asterisk. Residues 302, 303 and 304 correspond to the modified Ala-Pro-Val inhibitor of α LP³⁷.

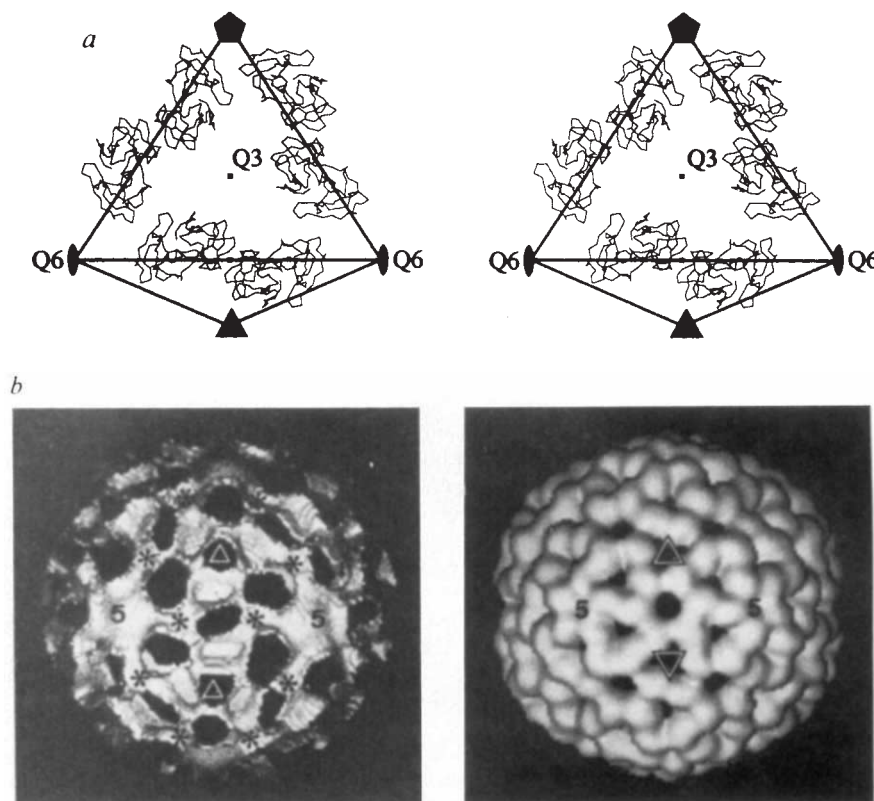


FIG. 6 Arrangement of dimers in a $T=4$ surface lattice with a mean protein radius of 185 Å. *a*, Stereographic view of three dimers related by a quasi-3-fold axis. Icosahedral 5-, 3- and 2-fold axes are shown with symbols. Quasi-6- and 3-fold axes are labelled with a Q. *b*, Comparison at 35 Å resolution of (left) electron-microscopy reconstruction of Sindbis cores³ with (right) reconstruction of a $T=4$ lattice made of SCP dimers as shown in *a*.

associated with the quasi-3-fold axes of a $T=4$ particle or with the quasi-2-fold axes of a $T=3$ particle^{3,41} (Fig. 1*b*). Thus, the $T=4$ surface lattice would require only one type of contact between the trimeric spikes and 3-fold sites on the core structure. Under certain conditions cores isolated from cytoplasm have an external radius of only 165 Å⁴³, instead of 205 Å when purified from infectious virions. The SCP dimers would be able to rotate about their 2-fold axes to make contact about the quasi-3-fold and icosahedral 3-fold axes. This, however, requires an appropriate radial shrinkage of the core particles, consistent with experimental observations.

It was also necessary to determine which side of the dimer goes to the outside and which to the inside of the core. Coombs and Brown⁷ used lactoperoxidase-mediated radioiodination of surface tyrosines of Sindbis cores to show that of the four tyrosines in SCP (162, 180, 189 and 198) only Tyr 180 was available on the core surface. The dimer structure shows that residue 180 is close to the surface of one side of the dimer whereas the other tyrosines are buried or on the opposite side of the dimer. With the dimer axis oriented appropriately, the amino end at residue Arg 114 would be pointing towards the outside of the core particle; but if the polypeptide were to make a short disordered connection to the uninterpreted density, then the basic end of the polypeptide would be directed towards the inside of the core.

About 2 and 31 of the C-terminal residues of E₁ and E₂, respectively, penetrate the viral membrane and interact with the nucleocapsid core^{22,44,45}. Knowledge of the core structure permits identification of the surface residues likely to make contact with the glycoprotein spikes. Comparison of alphavirus sequences⁴⁶ suggests that Met 137, Asn 172, Gly 201, Leu 231, Ala 234, Thr 238 and Ile 254 are involved in this interaction. These residues form a roughly linear trace, about 45 Å long, on the monomer surface near the 3-fold or quasi-3-fold axes. The surface residues of the core should also bind to a specific RNA sequence in the nonstructural protein-1 gene⁴⁷ and to the large subunit of ribosomes⁴⁸, properties that may be important in

assembly and disassembly as well as in virus-induced inhibition of host cell protein synthesis⁴⁹.

Evolution

Infectious single-stranded RNA virus genomes evolve mainly by divergence and by recombination⁵⁰. The core protein of Sindbis virus has a structural role which in many respects is analogous to that of the coat proteins of viruses. Nearly all other viruses, whether encapsidating RNA or DNA, use an eight-stranded β -barrel motif in the structure of the coat protein^{51,52,65}, although the icosahedral phage MS2⁵³ and the rod-shaped plant tobacco mosaic virus⁵⁴ are exceptions. Elucidation of the SCP structure shows that a member of the chymotrypsin-like serine proteinase family forms the coat of the internal structural core of an alphavirus. A common evolutionary origin has been suggested for other parts of the Sindbis virus genome, for tripartite plant viruses (such as alfalfa mosaic virus and bromo mosaic virus) and for tobacco mosaic virus^{55,56}. Hence, different components of the viral genome have apparently evolved from different sources. The SCP structure has identified a new evolutionary source for a viral core or coat protein. At this time alphaviruses are probably the only viruses where the proteinase activity of the capsid protein has been clearly established, and it will be of interest to determine whether other viruses contain similar molecules as structural proteins.

Most viruses that have single-stranded RNA genomes code for proteinases that are members of the chymotrypsin family⁵⁷. Animal picornavirus and plant como- and potyvirus polyproteins may also be related to serine proteinases, although the essential nucleophile is a cysteine^{58,59}. Further confirmation of the proposed homology is provided by the SCP structure determination. Figure 5 shows the proposed alignment⁵⁹ of the polio 3C proteinase to SCP, chymotrypsin and α -lytic proteinase. The position of deletions in SCP relative to chymotrypsin are closely mimicked by polio 3C proteinase, suggesting a more recent divergence between viral proteinases than between viral and other chymotrypsin-like proteinases. These similarities provide

some confirmation of the proposal that the viral cysteine proteinases are folded similarly to chymotrypsin. (The proposed catalytic triad of picornavirus and comovirus cysteine proteinases, based on their similarity to chymotrypsin, has recently been established by site-directed mutagenesis⁶⁶⁻⁶⁸.) Hence, as both SCP⁴⁷ and polio 3C proteinase⁶⁰ bind to specific RNA sequences, they may also have equivalent conserved RNA bind-

ing sites. Andino *et al.* (R. Andino, G. E. Rieckhof and D. Baltimore, personal communication) have shown, using site-specific mutagenesis, that polio RNA binds in the vicinity of the domain connection. The equivalent site on SCP would be on the outside of the core particle, consistent with the ability of the cores to bind to ribosomal RNA. □

Received 2 August; accepted 3 September 1991.

- Shope, R. E. in *The Togaviruses: Biology, Structure, Replication* (ed. Schlesinger, R. W.) 47-82 (Academic Press, New York, 1980).
- Freeman, R. & Leonard, K. R. *J. Microsc.* **122**, 275-286 (1981).
- Fuller, S. D. *Cell* **48**, 923-934 (1987).
- Horzinek, M. & Mussgay, M. *J. Virol.* **4**, 514-520 (1969).
- Enzmann, P. J. & Weiland, F. *Virology*, **95**, 501-510 (1979).
- Fuller, S. D. & Argos, P. *EMBO J.* **6**, 1099-1105 (1987).
- Coombs, K. & Brown, D. T. *Virus Res.* **7**, 131-149 (1987).
- Coombs, K. & Brown, D. T. *J. molec. Biol.* **195**, 359-371 (1987).
- Wengler, G., Boege, U., Wengler, G., Bischoff, H. & Wahn, K. *Virology*, **118**, 401-410 (1982).
- Harrison, S. C. in *The Togaviridae and Flaviviridae* (eds Schlesinger, S. & Schlesinger, M. J.) 21-34 (Plenum, New York, 1986).
- von Bonsdorff, C. H. & Harrison, S. C. *J. Virol.* **16**, 141-145 (1975).
- Kääriäinen, L. & Söderlund, H. *Curr. Topics Microbiol. Immunol.* **82**, 15-69 (1978).
- Simons, K. & Warren, G. *Adv. Protein Chem.* **36**, 79-132 (1984).
- Caspar, D. L. D. & Klug, A. *Cold Spring Harbor Symp. quant. Biol.* **27**, 1-24 (1962).
- Strauss, E. G. & Strauss, J. H. in *The Togaviridae and Flaviviridae* (eds Schlesinger, S. & Schlesinger, M. J.) 35-90 (Plenum, New York, 1986).
- Strauss, E. G., Rice, C. M. & Strauss, J. H. *Virology*, **133**, 92-110 (1984).
- Boege, U., Wengler, G. & Wittmann-Liebold, B. *Eur. J. Biochem.* **133**, 415-426 (1983).
- Strauss, J. H. & Strauss, E. G. *Semin. Virol.* **1**, 347-356 (1990).
- Simmons, D. T. & Strauss, J. H. *J. molec. Biol.* **86**, 397-409 (1974).
- Scupham, R. K., Jones, K. J., Sagik, B. P. & Bose, H. R. Jr *J. Virol.* **22**, 568-571 (1977).
- Aliperti, G. & Schlesinger, M. J. *Virology*, **90**, 366-369 (1978).
- Schlesinger, M. J. & Schlesinger, S. in *The Togaviridae and Flaviviridae* (eds Schlesinger, S. & Schlesinger, M. J.) 121-148 (Plenum, New York, 1986).
- Scheefers, H., Scheefers-Borchel, U., Edwards, J. & Brown, D. T. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7277-7281 (1980).
- Boege, U., Wengler, G. & Wittmann-Liebold, B. *Virology*, **113**, 293-303 (1981).
- Hahn, C. S., Strauss, E. G. & Strauss, J. H. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4648-4652 (1985).
- Melancon, P. & Garoff, H. *J. Virol.* **61**, 1301-1309 (1987).
- Hahn, C. S. & Strauss, J. H. *J. Virol.* **64**, 3069-3073 (1990).
- Boege, U. *et al. J. molec. Biol.* **208**, 79-82 (1989).
- Matthews, B. W. *J. molec. Biol.* **33**, 491-497 (1968).
- Kabsch, W. *J. appl. Cryst.* **21**, 916-924 (1988).
- Rossmann, M. G. & Blow, D. M. *Acta crystallogr.* **15**, 24-31 (1962).
- Harrison, S. C., Olson, A. J., Schutt, C. E., Winkler, F. K. & Bricogne, G. *Nature* **276**, 368-373 (1978).
- Abad-Zapatero, C. *et al. Nature* **286**, 33-39 (1980).
- Coombs, K. M. & Brown, D. T. *J. Virol.* **63**, 883-891 (1989).
- Birkhoff, J. J. & Blow, D. M. *J. molec. Biol.* **68**, 187-240 (1972).
- Fujinaga, M., Delbaere, L. T. J., Brayer, G. D. & James, M. N. G. *J. molec. Biol.* **184**, 479-502 (1985).
- Bone, R., Shenvi, A. B., Kettner, C. A. & Agard, D. A. *Biochemistry*, **26**, 7609-7614 (1987).

- Rossmann, M. G. & Argos, P. *J. biol. Chem.* **250**, 7525-7532 (1975).
- Bode, W., Schwager, P. & Huber, R. *J. molec. Biol.* **118**, 99-112 (1978).
- Wengler, G., Wengler, G., Boege, U. & Wahn, K. *Virology* **132**, 401-412 (1984).
- Vogel, R. H., Provencher, S. W., von Bonsdorff, C. H., Adrian, M. & Dubochet, J. *Nature* **320**, 533-535 (1986).
- Anthony, R. P. & Brown, D. T. *J. Virol.* **65**, 1187-1194 (1991).
- Coombs, K., Brown, B. & Brown, D. T. *Virus Res.* **1**, 297-302 (1984).
- Vaux, D. J. T., Helenius, A. & Mellman, I. *Nature* **336**, 36-42 (1988).
- Metsikkö, K. & Garoff, H. *J. Virol.* **64**, 4678-4683 (1990).
- Hahn, C. S., Lustig, S., Strauss, E. G. & Strauss, J. H. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5997-6001 (1988).
- Weis, B., Nitschko, H., Ghattas, I., Wright, R. & Schlesinger, S. *J. Virol.* **63**, 5310-5318 (1989).
- Ulmann, I., Söderlund, H. & Kääriäinen, L. *Virology* **99**, 265-276 (1979).
- Van Steeg, H., Kasperaitis, M., Voorma, H. O. & Benne, R. *Eur. J. Biochem.* **138**, 473-478 (1984).
- Strauss, J. H. & Strauss, E. G. *A. Rev. Microbiol.* **42**, 657-683 (1988).
- Rossmann, M. G. & Johnson, J. E. *A. Rev. Biochem.* **58**, 533-573 (1989).
- Tsao, J. *et al. Science* **251**, 1456-1464 (1990).
- Valegård, K., Liljas, L., Fridborg, K. & Unge, T. *Nature* **345**, 36-41 (1990).
- Namba, K. & Stubbs, G. *Science* **231**, 1401-1406 (1986).
- Haseleiff, J., Goeltz, P., Zimmern, D., Ahlquist, P., Dasgupta, R. & Kaesberg, P. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4358-4362 (1984).
- Ahlquist, P. *et al. J. Virol.* **53**, 536-542 (1985).
- Kräusslich, H. G. & Wimmer, E. *A. Rev. Biochem.* **57**, 701-754 (1988).
- Bazan, J. F. & Flettprick, R. *Proc. natn. Acad. Sci. U.S.A.* **85**, 7872-7876 (1988).
- Gorbalenya, A. E., Donchenko, A. P., Blinov, V. M. & Koonin, E. V. *FEBS Lett.* **243**, 103-114 (1989).
- Andino, R., Rieckhof, G. E. & Baltimore, D. *Cell* **63**, 369-380 (1990).
- Kraulis, P. *J. appl. Cryst.* (in the press).
- Connolly, M. L. *J. molec. Graphics* **3**, 19-24 (1985).
- Liljas, A. & Rossmann, M. G. *A. Rev. Biochem.* **43**, 475-507 (1974).
- Rossmann, M. G. *et al. J. appl. Cryst.* (in the press).
- McKenna, R. *et al. Nature* (submitted).
- Kean, K. M., Teterina, N. L., Marc, D. & Girard, M. *Virology* **181**, 609-619 (1991).
- Hämmerle, T., Hellen, C. U. T. & Wimmer, E. *J. biol. Chem.* **266**, 5412-5416 (1991).
- Dessens, J. T. & Lomonosoff, G. P. *Virology* **184**, 738-746 (1991).

ACKNOWLEDGEMENTS. H.K.C. and L.T. are equal first authors. We thank T. J. Smith and J. Tsao for their advice and participation in the initial stages of this project, and M. Luo, A. Brunger, D. Brown, M. James, M. Laskowski, R. Andino and J. Strauss for discussions. We thank S. Fuller, H. Cheng and T. S. Baker for assistance in the preparation of Fig. 6b, E. Westbrook and M. Deacon (Argonne National Laboratory) and N. Jones and R. Schevitz (Eli Lilly Company) for use of their area detectors, and H. Prongay and S. Weller for help in preparing the manuscript. The work was supported by grants from the NSF and the NIH to M.G.R. and by the Deutsche Forschungsgemeinschaft to G.W. Additional support for structural studies came from an NIH AIDS Center grant and the Lucille P. Markey Foundation award.

LETTERS TO NATURE

Origin of the Napoleon's hat nebula around SN1987A and implications for the progenitor

Ph. Podsiadlowski, A. C. Fabian & I. R. Stevens

Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK

THE emission nebula around supernova 1987A^{1,2} is recombination radiation excited by the ultraviolet flash from the supernova explosion as it travels through circumstellar material. Recent observations show it to have a complex but clearly structured morphology, which has been likened to Napoleon's hat³. We present here a simple geometrical model for the nebula, consisting of a ring and a truncated double cone. When the effects of light travel-time are included, the model reproduces the important topological structures of the nebula and makes detailed quantitative predictions for its future appearance. In particular, the hat-shaped northern rim is simply explained as the interaction of the light front with the northern cone. To explain the origin of the double cone, we argue that the progenitor of SN1987A was in a binary system: its strong wind, colliding with a weaker wind from the companion star, created an asymptotic shock surface that was spread out into the required geometry by the rotation of the binary.

This suggestion is consistent with other binary models¹³ for the progenitor of this unusual supernova.

As all the structures around SN1987A (see Fig. 1a) were generated before the explosion by the star or stars in the immediate neighbourhood of the supernova, they provide a direct imprint of the progenitor's eventual past. Understanding the origin of the structures will therefore lead to a better understanding of the progenitor's evolution and may help to explain the anomalous features of this supernova event.

In Fig. 2a, we present a simple geometrical model for the nebula. In the model, the dense, circumstellar material (outside the central ring nebula²) is distributed on the surface of a truncated, elliptical double cone, centred on the supernova. The cone's opening angle with respect to its symmetry axis varies between 40° and 60°, and the cone is truncated at a distance of 2.5'' from the supernova (which corresponds to a distance of 1.9×10^{18} cm for a supernova distance of 1.7×10^5 light years³). The cone's precise orientation with respect to the central ring and the observer is shown in Fig. 2a. To illustrate the optical appearance of the nebula projected onto the plane of the sky (see Fig. 1b), we adopted a simple, not necessarily realistic model for the nebula's emission. The only important structure that (weakly) depends on the emission model is the lower part of the hat structure.

Our model reproduces the morphology of the nebula in detail. The northern rim of the hat structure and, in particular, the

curvature change along the rim are well reproduced. Indeed, the three-dimensional space position of the rim is uniquely determined by the model-independent assumption that it represents the intersection of the nebula with the illumination front (which bounds the glowing portions of the nebula and which, for a distant observer, has the geometry of a paraboloid; see Fig. 2a). The excellent agreement of our model therefore implies that, in any model, the northern rim must lie on an elliptical cone. The model also predicts a moving light echo coincident with the northern rim, since the rim marks the leading edge of the illumination front. Such a light echo has indeed been observed⁴, further strengthening the model. If this fact is combined with the available velocity information for the inner loops¹, which implies that the northern loop is in front of the supernova, whereas the southern loop is behind it, it follows immediately that the geometry of a truncated, eccentric double cone is the simplest possible geometry that can connect the hat structure with the inner loops.

The dense rings, which form the ends of the double cone near the supernova, consist, in the model discussed later, of material that originally was part of the double cone, but was swept up by the energetic wind generated in the progenitor's final blue-

supergiant phase. In our emission model, these rings do not appear as uniform, closed loops. The reason is that emission from matter on the southern part of the northern ring, which was excited by the ultraviolet flash soon after the supernova and which had time to recombine, has already become relatively faint, whereas the northern part of the southern ring has not yet been excited by the ultraviolet flash (see Fig. 2a). This may explain why Wampler *et al.*¹ interpret these structures as filamentary loops rather than as closed rings. We note, however, that this interpretation depends, unlike most of our other conclusions, on the adopted emission model. Furthermore, we have made no attempt to model another filament (east of the supernova), which we attribute to an interaction of the progenitor's wind with a second star.

The apparent ellipticity of these rings is somewhat larger in our model than is observed. This small discrepancy is probably caused by our assumption that the double cone is truncated at a fixed distance from the supernova, whereas this truncation is not necessarily expected in the cone's formation scenario discussed below. We also note that the disagreement is largest east of the supernova, where the cone structure is likely to be distorted by the presence of the second star.

The geometrical model presented here makes very definite, quantitative predictions, because, if combined with a realistic emission model, it completely determines the future appearance

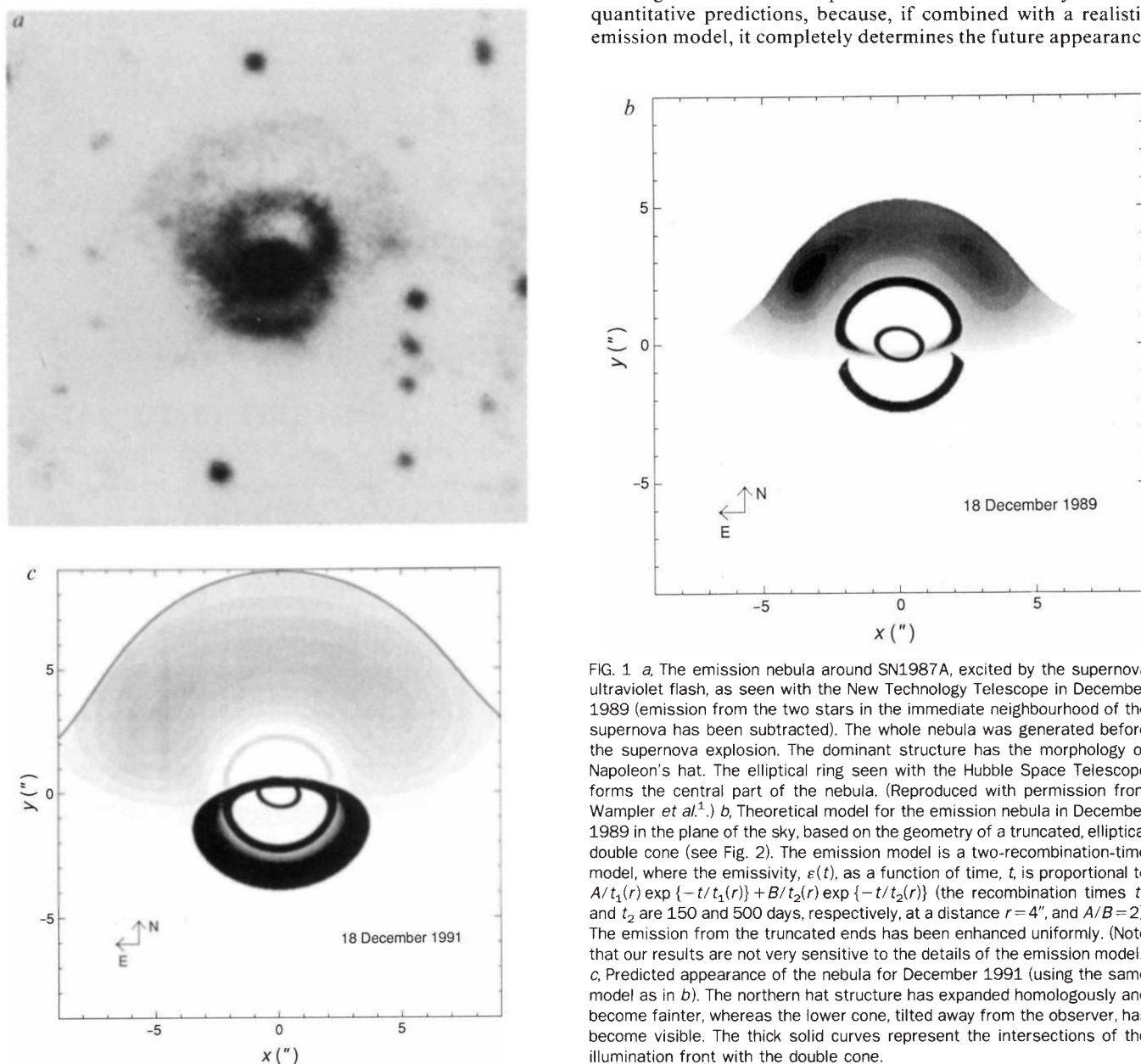


FIG. 1 *a*, The emission nebula around SN1987A, excited by the supernova ultraviolet flash, as seen with the New Technology Telescope in December 1989 (emission from the two stars in the immediate neighbourhood of the supernova has been subtracted). The whole nebula was generated before the supernova explosion. The dominant structure has the morphology of Napoleon's hat. The elliptical ring seen with the Hubble Space Telescope forms the central part of the nebula. (Reproduced with permission from Wampler *et al.*¹.) *b*, Theoretical model for the emission nebula in December 1989 in the plane of the sky, based on the geometry of a truncated, elliptical double cone (see Fig. 2). The emission model is a two-recombination-time model, where the emissivity, $\epsilon(t)$, as a function of time, t , is proportional to $A/t_1(r) \exp\{-t/t_1(r)\} + B/t_2(r) \exp\{-t/t_2(r)\}$ (the recombination times t_1 and t_2 are 150 and 500 days, respectively, at a distance $r=4''$, and $A/B=2$). The emission from the truncated ends has been enhanced uniformly. (Note that our results are not very sensitive to the details of the emission model.) *c*, Predicted appearance of the nebula for December 1991 (using the same model as in *b*). The northern hat structure has expanded homologously and become fainter, whereas the lower cone, tilted away from the observer, has become visible. The thick solid curves represent the intersections of the illumination front with the double cone.

of the nebula. In Fig. 1c, we illustrate the appearance of the nebula for December 1991. We emphasize that, because of the simplicity of our emission model, we can reliably predict only morphological features (such as the location of the illumination fronts), but not the relative emissivity across the nebula. The northern hat structure has expanded homologously (the top of the hat expands at a superluminal rate of $\sim 1.9''$ per yr) and has started to fade away, whereas the southern cone, which is tilted away from the observer, slowly becomes visible. The glowing portion of the southern cone will develop the shape of a crescent, and its southern edge will expand at a rate of $\sim 0.8''$ per yr. (The latter prediction is less secure, as our geometrical model is only weakly constrained by the southern cone, and different orientations for the southern cone may yield similarly good fits.)

Although the double-cone model makes straightforward predictions, it is not immediately obvious how this structure would have formed. Different formation methods can be imagined (for example, some bipolar flows⁵ may be able to generate very similar structures). We suggest here a different possibility, namely that the double cone represents the asymptotic shock surface of two colliding winds in a massive, close binary (for a detailed discussion of the colliding wind phenomenon, see ref. 6). In this model, the precursor to the supernova system would previously have been a close binary, resembling the system of ζ Aurigae⁷ or VV Cephei⁸; that is, a binary with an orbital period of several years, consisting of a red supergiant (with a mass of $\sim 15 M_{\odot}$) and an early-type companion (with a mass of ~ 5 – $10 M_{\odot}$). To see how the collision of the winds from these components can produce conical gas structures, we first consider such a collision in two dimensions and ignore binary motion. Then, the shape of the shock structure can be found by consider-

ation of the ram-pressure balance⁶. The wind with greater momentum flux (from the red supergiant, star A) will tend to dominate the weaker wind (from the companion, star B), forming a shock with a hyperboloidal surface around star B. From the perspective of star A, the great extent of the shock structure lies at an asymptotic angle, α_s , from the line joining the centres of the system⁹. For typical wind parameters (see, for example ref. 10), the momentum ratio of the two winds is ~ 10 ; this gives rise to a value of $\alpha_s \approx 45^\circ$ (ref. 8).

The situation becomes considerably more complicated when binary orbital motion becomes important, as then one must deal with the three-dimensional structure of this shock surface. The most important effect of the orbital motion is that the shock surface around star B will rotate around the orbital axis. When the orbital velocity is faster than the velocity of the red supergiant wind (as is the case here), the orbital motion will wind the shock structure around star B. Velocity differences (in the winds from star A and B and in the post-shock flow) will cause further shocks in the material in this structure, which therefore forms a single surface axisymmetric about the orbital axis of the binary. The asymptotic opening angle of the structure remains $\sim \alpha_s$, so at large radial distance the structure resembles a double cone of opening angle $90^\circ - \alpha_s$ (note that the opening angle changes with time, as the relative momentum ratio of the winds evolves). The shock surface is shown schematically in Fig. 2b. Although detailed verification must await a proper three-dimensional hydrodynamical treatment of the colliding winds, it is worth noting that binary systems in which colliding winds are important often show similar or similarly complex structures (see for example refs 11, 12). High-resolution radio observations of ζ Aurigae or VV Cephei systems could also provide

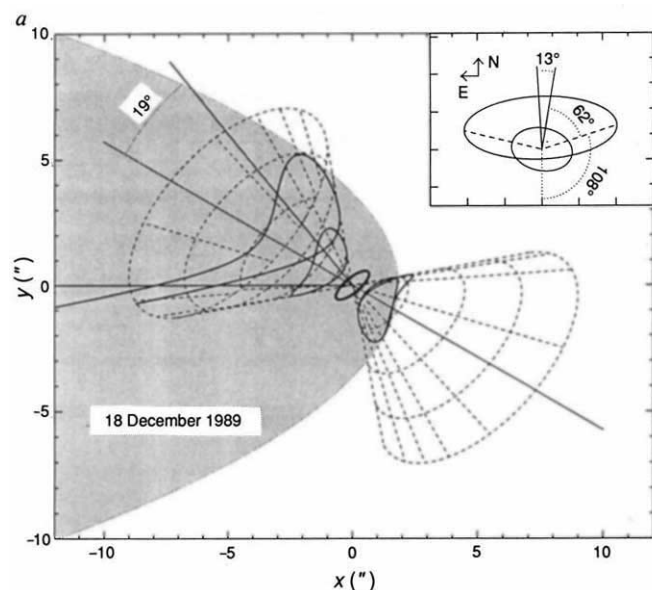
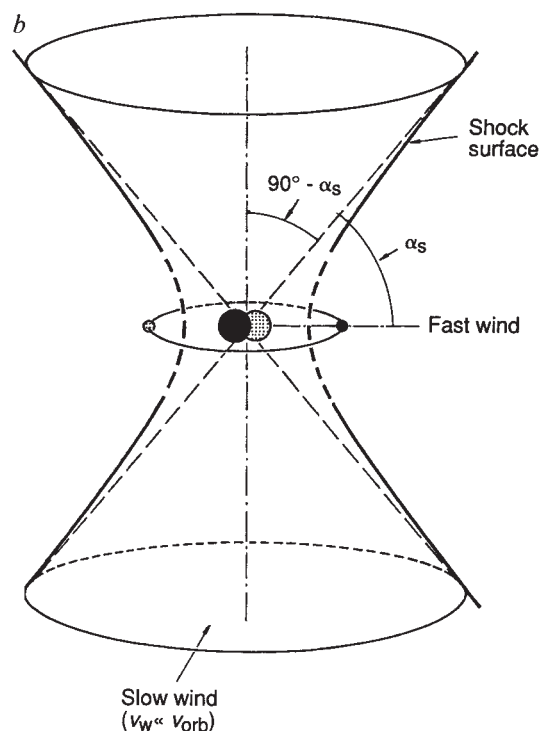


FIG. 2 a, Three-dimensional view of the geometry and orientation of the double-cone model. The dashed curves outline the bi-conical structure. The thin, solid straight lines represent (clockwise from 0°) the direction of the observer, the symmetry axis of the double cone and the rotation axis of the central ring. The thick solid curves mark the truncated ends of the cone (at a distance of $2.5''$ from the supernova), the intersections of the illumination front with the cone and the central ring. The shaded area indicates the portions of the nebula that have been excited by the supernova ultraviolet flash by December 1989, as seen from an observer on Earth (who is to the left, as indicated). Inset: schematic view of the orientation of the northern cone (only one elliptical cross-section is shown) and the central ring in the plane of the sky. The solid straight lines show (from left to right) the projected orientation of the cone's symmetry axis and the central ring's rotation axis. The dashed curves mark the largest opening of the cone with respect to its symmetry axis (the cone's opening angle varies from 40° to



60°). The parameters for the central ring are taken from ref. 2. b, A schematic picture of a colliding wind scenario for the origin of the biconical nebula. The slow, but dominant red-supergiant wind collides with a weaker wind from its companion. When the velocity of the red-supergiant wind is less than the orbital velocity of the binary, the shock surface is wound into an axisymmetric structure about the orbital axis of the binary (solid curves). This structure asymptotically approaches a double cone. We note that the geometry of the shocked surface near the orbital plane is likely to be more complicated.

some independent observational evidence for it, as they might reveal the presence of bi-conical shock structures in these systems.

The origin of the eccentricity of the cone is not immediately obvious. One possibility is that the binary orbit was eccentric—systems of this type often have large eccentricities^{7,8}. Other possibilities involve interactions with the interstellar medium or with large-scale magnetic fields.

A colliding wind model is consistent with previously published binary models for SN1987A¹³. For example, in a model involving the merging of the system^{14–16}, the double-cone structure may have been produced while the primary was ascending the asymptotic-giant branch. The system then merged ($\sim 10^4$ – 10^5 yr before the supernova explosion), forming in the process a rapidly rotating single star with a thoroughly mixed envelope. While the star shrank to become a blue supergiant, the envelope was spun up to near break-up velocity, and substantial mass was lost in the equatorial plane (forming an excretion disk around the system). In the final blue-supergiant phase, the energetic

blue-supergiant wind swept up all the structures that had been produced previously; in particular, it swept the excretion disk into a ring, and the conical structures into a truncated double cone. This would have produced precisely the structures of our model. Such a model could potentially explain all the chief anomalies of the supernova (the blue colour of the progenitor, the chemical anomalies in the progenitor's envelope^{17,18} and the surrounding structures; see Ph.P., Institute of Astronomy, preprint). Future observations of the nebula will be important, not only to help verify our model for the nebula, but also to provide an important step towards understanding this unusual supernova event.

Since, if a merger had occurred, the immediate supernova progenitor would have been a single blue supergiant without a surviving companion, direct verification of a merger model would be difficult. The nebula around the supernova has now provided us with a smoking gun, which may lead to a confirmation of such a model; in the long run, this may be almost as conclusive as a surviving companion. $\square$

Received 6 May; accepted 9 September 1991.

1. Wampler, E. J. *et al. Astrophys. J.* **362**, L13–L16 (1990).
2. Jacobson, P. *et al. Astrophys. J.* **369**, L63–L66 (1991).
3. Schwarzschild, B. *Physics Today* **44** (No. 2), 20 (1991).
4. Crots, A. P. S. & Kunkel, W. E. *Astrophys. J.* **366**, L73–L77 (1991).
5. Morris, M. in *From Miras to Planetary Nebulae: Which Path for Stellar Evolution?* (eds Mennessier, M. O. & Omont, A.) 520–535 (Editions Frontiers, Gif-sur-Yvette 1990).
6. Luo, D., McCray, R. & Mac Low, M.-M. *Astrophys. J.* **362**, 267–273 (1990).
7. Lee, E. K. & Wright, K. O. *Publ. Dominion Astrophys. Obs.* **11**, 339–361 (1960).
8. Wright, K. O. *J. R. astr. Soc. Canada* **72**, 152–193 (1977).
9. Girard, T. & Willson, L. A. *Astr. Astrophys.* **183**, 247–256 (1987).
10. Chiosi, C. & Maeder, A. A. *Rev. Astr. Astrophys.* **24**, 329–375 (1986).

11. White, R. L. & Becker, R. H. *Astrophys. J.* **297**, 677–685 (1985).
12. Kenny, H. T., Taylor, A. R. & Seaquist, E. R. *Astrophys. J.* **366**, 549–559 (1991).
13. Hsu, J. J. L., Joss, P. C., Podsiadlowski, Ph. & Rappaport, S. in *Supernovae* (ed. Woosley, S.) 191–197 (Springer, New York, 1990).
14. Chevalier, R. A. & Soker, N. *Astrophys. J.* **341**, 867–882 (1989).
15. Hillebrandt, W. & Meyer, F. *Astr. Astrophys.* **219**, L3–L6 (1989).
16. Podsiadlowski, Ph., Joss, P. C. & Rappaport, S. *Astr. Astrophys.* **227**, L9–L12 (1990).
17. Williams, R. E. *Astrophys. J.* **320**, L117–L120 (1987).
18. Höflich, P. *Proc. astr. Soc. Austr.* **7**, 434–442 (1990).

ACKNOWLEDGEMENTS. We thank E. J. Wampler for his permission to reproduce one of his images of the supernova nebula.

High-resolution VLA images of the Galactic Centre at 2-cm wavelength with large dynamic range

Jun-Hui Zhao*, W. M. Goss*, K. Y. Lo† & R. D. Ekers‡

* National Radio Astronomy Observatory, Socorro, New Mexico 87801, USA

† Astronomy Department, University of Illinois, 1002 W. Green Street, Urbana, Illinois 61801, USA

‡ Australia Telescope National Facility, PO Box 76, Epping, New South Wales 2121, Australia

THE existence of a black hole at the centre of our galaxy was proposed by Lynden-Bell and Rees¹, and the later discovery of a bright compact radio source (Sgr A*) at the Galactic Centre² suggested the presence of a central exciting source³. Recently, Yusef-Zadeh *et al.*⁴ reported faint extended radio structures (the 'satellite radio features') which they interpreted as thermally emitting blobs of plasma physically associated with Sgr A*. These puzzling features could be taken as evidence for an outflowing wind from Sgr A*/IRS16 (a complex infrared source), but there was also a suspicion that they might be instrumental artefacts. We have re-observed the Galactic Centre using the Very Large Array, obtaining images at a wavelength of 2 cm with resolution of ~ 0.1 arcsec (0.005 pc). The r.m.s. noise of 70 μ Jy per beam, with a dynamic range of 22,000:1, is comparable to that of ref. 4, but the new structures reported in ref. 4 do not appear in our observations. We discuss several ways that low-level artefacts near Sgr A* might have been generated. We also confirm the presence of a plume of emission 1 arcsec southwest of Sgr A*.

Early observations with the Very Large Array (VLA)^{5–7} revealed that much of the radio continuum free-free emission in the central few parsecs of the Galaxy (Sgr A West) is distributed in a spiral-shaped filamentary structure surrounding

Sgr A*. The kinematics of the ionized gas responsible for the radio continuum emission have been determined from radio recombination lines^{8,9} and [Ne II]¹⁰ observations. More recent high-resolution 2-cm VLA observations¹¹ have revealed faint details in the radio emission, such as a small cavity of size $\sim 2''$ located $\sim 3''$ southwest of Sgr A*, and the radio counterpart of IRS7, a supergiant star located $\sim 5''$ north of Sgr A*. Such features, along with the 'satellite features' (ref. 4), have been taken to be evidence for the existence of stellar winds¹¹ and a hot, high-velocity wind originating from Sgr A* itself¹².

High-resolution radio imaging of faint structures near Sgr A* is difficult for three reasons: (1) Sgr A* (~ 1 Jy) is a time-variable source; (2) the source is located in the middle of a strong extended source (Sgr A West) of ~ 20 Jy at 2 cm; (3) weak artefacts are more likely to arise in regions of high sidelobes near a point source. Because of these difficulties, various faint artefacts can be produced close to a strong point source such as Sgr A*. The faint structures (features α to δ) near Sgr A* observed in ref. 4, which are near Sgr A* and which are only $\sim 0.1\%$ of the peak intensity, could be such artefacts, as the authors suggested.

During 1990, we observed the Galactic Centre using the VLA in the A, B and C configurations at 2 cm. As the primary goal of these observations was to monitor the flux density variations of Sgr A*, the observations were performed in the snapshot mode with total observation times of 5–20 min on source. The total number of observations was ~ 20 , separated by intervals of typically 2 weeks. The visibility data were initially calibrated in amplitude and phase using 3C286 and 1748–253, assuming that the flux density of 3C286 was 3.44 at 2 cm. Before combining the data, the individual snapshot databases were self-calibrated¹³ in phase for two or three iterations. As it was known that Sgr A* varies substantially in flux density ($\sim 50\%$ within three months)¹⁴, fluctuations in the combined visibility data caused by the time-variable amplitude of the point source could produce various artefacts in the images.

The varying amplitude scales in different observations due to the point source must be corrected before combination. The

best approach is to subtract or add flux density corresponding to a point source at the position of Sgr A* in the Fourier domain (known as the visibility ($u-v$) domain) to scale the point source to a constant flux density. The advantage of this approach (instead of subtracting the point source from all the observations) is that a point-source model can still be used in further self-calibrations for the combined data set. In addition, there could be discrepancies of reference positions in the different observations because of problems such as a phase gradient across the array caused by the troposphere at this low elevation. The reference positions in each snapshot database were aligned to a common position by using the point source before the combination procedure. The common flux-density scale and reference position are necessary to achieve high-dynamic-range images. New gain solutions can then be calculated by comparing the visibility data separately with a single model for all the observations in the self-calibration procedure. The remaining antenna-based phase and amplitude errors can be effectively minimized in this way.

The 'dirty images' were then deconvolved with the dirty interferometer beam using both the CLEAN¹⁵ algorithm and the maximum-entropy method¹⁵ (MEM). All of the images presented here were produced using MEM following an initial subtraction of essentially all of the flux density of a point source (from Sgr A*) from the visibility data. The final r.m.s. noise for each image is within a factor of two of the theoretical thermal noise. To test the fidelity of the images, the visibility data at 2 cm have been divided into two different databases with independent visibility sampling coverage; the images made from independent processing (self-calibration and deconvolution) agree with each other at a level of 4σ .

MEM produces two types of images: (1) the pixel image, in which the intensity is measured in Jy per pixel and the spatial resolution is signal-dependent and somewhat higher than the corresponding CLEAN map, and (2) the pixel image convolved

with a CLEAN beam, in which residuals are added back to form a 'restored' image.

To compare our results directly with those in ref. 4, we present a contour map of the 2-cm radio emission from the region surrounding Sgr A* (Fig. 1a), made from a MEM pixel image using the same contour levels and pixel size as the authors used in their Fig. 1a (a sketch of the earlier results is given for comparison in Fig. 1b). The restored MEM map of the same region is presented in Fig. 1c. The full width at half maximum of the relevant CLEAN beam is $0.22'' \times 0.12''$ (position angle -0.22°) with an r.m.s. of $70 \mu\text{Jy}$ per beam (dynamic range 22,000:1). Here, the r.m.s. noise is measured in the final restored MEM images and the peak intensity is rescaled to 1.6 Jy per beam, the highest peak intensity observed during the one-year period. For comparison, the image published by ref. 4 has an r.m.s. noise of $75 \mu\text{Jy}$ per beam. The extended features (α to δ) close to Sgr A* reported by ref. 4 would have typical intensities of 1×10^{-4} Jy per pixel in Fig. 1a and 1 mJy per beam in Fig. 1c, and would therefore be 10 times the r.m.s. noise, yet they were not detected.

The inner $15''$ of the Galactic Centre at 2 cm is presented in Fig. 2. The radio counterpart¹¹ of IRS7 can clearly be seen in the image $\sim 5''$ north of Sgr A*. A 'cavity' of size $2.5''$ (0.1 pc) appears $\sim 3''$ southwest of Sgr A*. The extended sources ϵ , ζ and η of ref. 4 close to and/or within this cavity are confirmed. The feature ϵ is marked in a corresponding grey-scale contour map in Fig. 2. The boundary of the cavity is indicated by a bold line in Fig. 2. In contrast to Yusef-Zadeh *et al.*¹¹, we find that there is a 'pocket' of size $1''$ southeast of the cavity (see Fig. 2). These two features appear to be connected.

A possible cause for the non-detection of the features (α to β) in 1990 could be a rapid expansion of the features since the earlier observations⁴ in 1986. The implied expansion velocity, however, is $\geq 10^3 \text{ km s}^{-1}$, which would indicate that the features were extremely hot ($T_e > 10^8 \text{ K}$). This explanation therefore

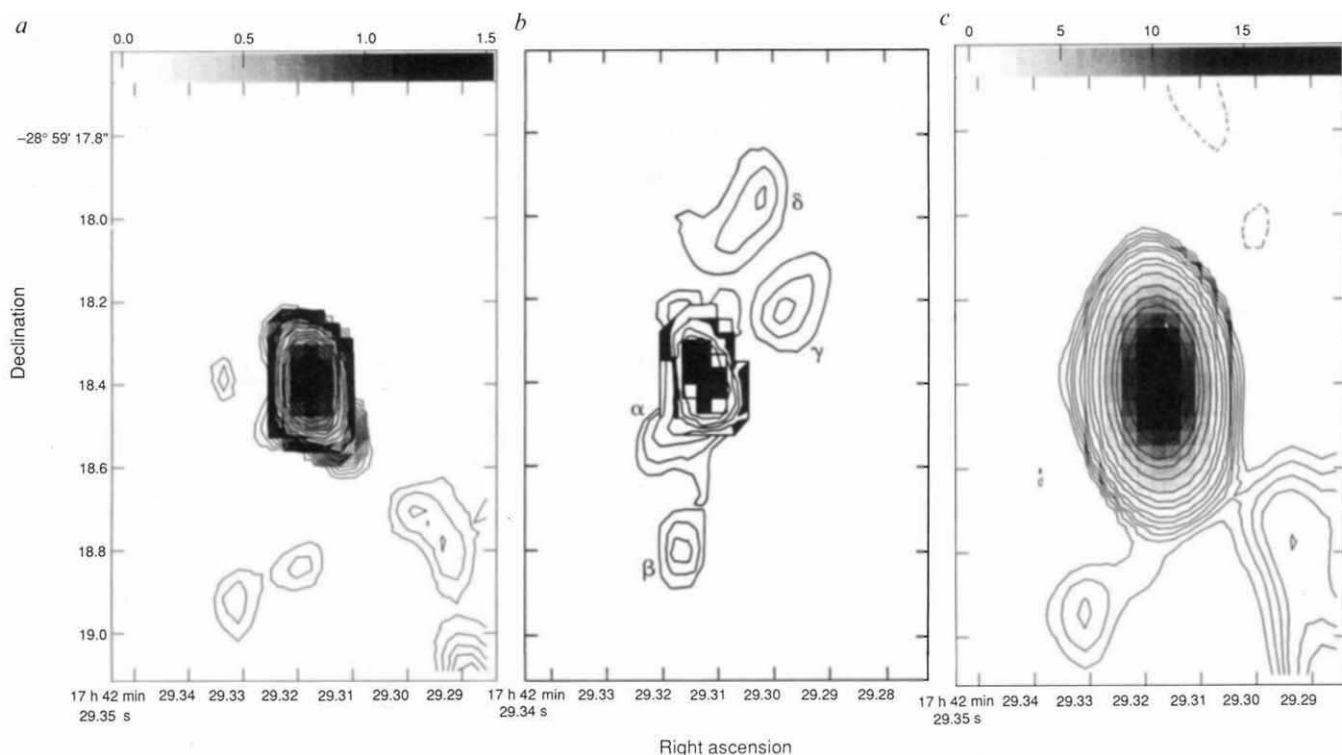


FIG. 1 Contour maps of Sgr A* made from MEM images at 2 cm with $0.035'' \times 0.035''$ cell size. *a*, MEM pixel image with super-spatial resolution $0.14'' \times 0.07''$ (P.A. = -0.22°); contours are 2, 3, 4, ..., 14, 16, 18, 20, 25, 30, 40, 50, 60, 80, 100, 130, 160×10^{-5} Jy per pixel. The grey-scale flux density range is 0 to 1.5 mJy per pixel. *b*, Sketch of Fig. 1a from Yusef-Zadeh *et al.*⁴, shown for comparison. The peak intensity of the features (α to δ) is

$\sim 1 \times 10^{-4}$ Jy per pixel, and the contours in the 'sketch' extend roughly to 5×10^{-5} Jy per pixel which is a factor of 2.5 higher than the minimum contours in *a*. *c*, Restored MEM image with a beam of $0.22'' \times 0.12''$ (P.A. = -0.22°); contours are $-1.5, 3, 4, 5, 6, 8, 10, 15, 20, 30, 50, 75, 100, 150, 200 \times \text{r.m.s.}$ (7×10^{-5} Jy per beam). The grey-scale flux-density range is -0.2 to 20 mJy per beam.

seems unlikely. Alternatively, the non-detection of the features might be attributed to the divergence of a rapid ejected gas. The kinetic energy required for the rapidly moving gas would be $>10^{46}$ erg. This energy is at least $100\times$ larger than the estimated total energy of Sgr A* (see, for example, ref. 16), suggesting that this explanation is also implausible.

To understand the possible origins of artefacts near a strong point source, we have carried out a number of experiments and found several processes that could cause such features ($\sim 0.1\%$ of the peak). (1) Artefacts can be produced by combinations of synthesis interferometer data arising from a point source with a time-variable flux density¹⁷. The errors due to a time-variable point source can be essentially eliminated by the method described above. (2) If the final model used in the self-calibration procedure contains significant phase and amplitude errors, faint artefacts near a point source can also be produced¹⁸. Some artefacts surrounding Sgr A* were observed in initial images produced from the current data; this was because errors in the visibility data were not effectively removed, as a model that was used in the self-calibration procedures included these errors. The intensities and the sizes of these features are similar to those reported in ref. 4. These phase and amplitude errors can be substantially removed by self-calibration with a more conservative model, for example by starting with a point-source model and an adequate visibility data restriction. The similarity of the features reported in ref. 4 to self-calibration artefacts makes the detection of such faint features very difficult in practice. (3) If a point source is not centred on a grid point, low-level artefacts close to the point source can also be produced. This occurs because a larger number of CLEAN components are needed to model an offset point source and these are more likely to include erroneous components. (4) Finally, the successful use of the MEM method requires that a point source should be completely subtracted before deconvolution¹⁵. Otherwise, faint, extended artefacts surrounding a point source are favoured by the maximum entropy solution¹⁹. In summary, we do not confirm the 'satellite radio features', labelled α , β , δ and γ , reported in ref. 4.

We do confirm a feature between Sgr A* and the 'cavity' to the southwest. This feature (labelled ϵ in ref. 4) has been observed previously at 2 and 6 cm²⁰ and is $\sim 1''$ from Sgr A*. In

Fig. 2, the feature appears to be a slightly elongated plume ($2'' \times 1.2''$) originating from Sgr A*. The integrated flux density of the feature at 2 cm is 35 mJy. The position of the emission plume is also near IRS16C and IRS16NW^{21,22}, but is not coincident with any of the IRS sources. The orientation of the elongated plume and the asymmetric distribution of the radiation intensity along the rim of the cavity (see Fig. 2) suggest that the cavity is a result of an impact of a wind on the ionized gas in Sgr A West. The large excess of radio emission in the western side of the cavity may indicate that the ionized gas has been compressed by the flow. It is possible that an outflow can bring sufficient mechanical energy to punch a hole in the ionized medium^{23,24}. It is unclear if the feature ϵ is physically associated with Sgr A* and/or IRS16. The spectral index of the radio emission is uncertain because of its faintness. We plan more VLA observations to investigate the physical connection between Sgr A*, feature ϵ and IRS16. $\square$

Received 18 June; accepted 30 August 1991.

1. Lynden-Bell, D. & Rees, M. J. *Mon. Not. R. astr. Soc.* **152**, 461–475 (1971).
2. Balick, B. & Brown, R. L. *Astrophys. J.* **194**, 265–270 (1974).
3. Lo, K. Y. *IAU Symp. No. 136, The Centre of the Galaxy* (ed. Morris, M.) 443–451. (Kluwer, Dordrecht, 1989).
4. Yusef-Zadeh, F., Morris, M. & Ekers, R. D. *Nature* **348**, 45–47 (1990).
5. Ekers, R. D., van Gorkom, J. H., Schwarz, U. J. & Goss, W. M. *Astr. Astrophys.* **122**, 143–150 (1983).
6. Brown, R. L., Johnston, K. J. & Lo, K. Y. *Astrophys. J.* **250**, 155–159 (1981).
7. Lo, K. Y. & Claussen, M. J. *Nature* **306**, 647–651 (1983).
8. Roberts, D. A., Goss, W. M., van Gorkom, J. H. & Leahy, J. P. *Astrophys. J.* **366**, L15–18 (1991).
9. van Gorkom, J. H., Schwarz, U. J. & Bregman, J. D. *Proc. IAU Symp. No. 106* (eds van Woerden, H., Burton, W. B. & Allen, R. J.) 371–376. (Reidel, Dordrecht, 1985).
10. Serabyn, E. & Lacy, J. H. *Astrophys. J.* **293**, 455–458 (1985).
11. Yusef-Zadeh, F., Morris, M. & Ekers, R. D. *IAU Symp. No. 136, The Centre of the Galaxy* (ed. Morris, M.) 443–451. (Kluwer, Dordrecht, 1989).
12. Yusef-Zadeh, F. & Morris, M. *Astrophys. J.* **371**, L63–67 (1991).
13. Cornwell, T. & Fomalont, E. B. *Synthesis Imaging in Radio Astronomy* (eds Perley, R. A., Schwab, F. R. & Bridle, A. H.) 185–196 (1989).
14. Zhao, J.-H., Goss, W. M., Lo, K. Y. & Ekers, R. D. *Relationships between Active Galactic Nuclei and Starburst Galaxies* (ed. Filippenko, A.) (Astronomical Society of the Pacific, San Francisco, in the press).
15. Cornwell, T. & Braun, R. *Synthesis Imaging in Radio Astronomy* (eds Perley, R. A., Schwab, F. R. & Bridle, A. H.) 167–184 (Astronomical Society of the Pacific, San Francisco, 1989).
16. Davies, R. D., Walsh, D. & Booth, R. S. *Mon. Not. R. astr. Soc.* **177**, 319–333 (1976).
17. Cotton, W. D. *Synthesis Imaging in Radio Astronomy* (eds Perley, R. A., Schwab, F. R. & Bridle, A. H.) 233–246 (Astronomical Society of the Pacific, San Francisco, 1989).
18. Ekers, R. D. *Synthesis Imaging in Radio Astronomy* (eds Perley, R. A., Schwab, F. R. & Bridle, A. H.) 99–212 (Astronomical Society of the Pacific, San Francisco, 1989).
19. Cornwell, T. J. & Evans, K. F. *Astr. Astrophys.* **143**, 77–83 (1985).
20. Killeen, N. E. B. & Lo, K. Y. *IAU Symp. No. 136, The Centre of the Galaxy* (ed. Morris, M.) 453–455 (Kluwer, Dordrecht, 1989).
21. Tollestrup, E. V., Carpp, R. W. & Becklin, E. E. *Astr. J.* **98**, 204–216 (1989).
22. DePoy, D. L. & Sharp, N. A. *Astr. J.* **101**, 1324–1328 (1991).
23. Castor, J., McCray, R. & Weaver, R. *Astrophys. J.* **200**, L107–110 (1975).
24. Weaver, R., McCray, R. & Castor, J. *Astrophys. J.* **218**, 377–395 (1977).

ACKNOWLEDGEMENTS. We thank T. Cornwell, F. Yusef-Zadeh, M. Morris and J. H. van Gorkom for comments. The NRAO is operated by Associated Universities, Inc., under a cooperative agreement with the NSF.

Origin of the shuttle glow

R. A. Viereck*, Edmond Murad*, B. David Green†, P. Joshi†, C. P. Pike*, R. Hieb‡ & G. Harbaugh‡

* Phillips Laboratory (WSSI), Hanscom AFB, Massachusetts 01731, USA

† Physical Sciences Inc., Andover, Massachusetts 01810, USA

‡ NASA, Johnson Space Center, Houston, Texas 77001, USA

A GLOW around exposed surfaces of the space shuttle facing the direction of orbital motion was first seen in 1983^{1,2}. This 'shuttle glow' extends about 10 cm from the surfaces, is peaked in wavelength at 680 nm, and within a resolution of about 3.5 nm forms a continuum^{3–6}. Similar anomalies were reported in rocket experiments as long ago as 1958⁷ and in more recent space-based studies⁸. Apart from its interest as an unusual physical phenomenon, shuttle glow may be a source of interference in space-based spectroscopy; anomalous airglow observations made by the Atmospheric Explorer spacecraft^{9,10} have been attributed to it. The most likely explanation seems to be the recombination of fast oxygen atoms in the upper atmosphere with NO absorbed

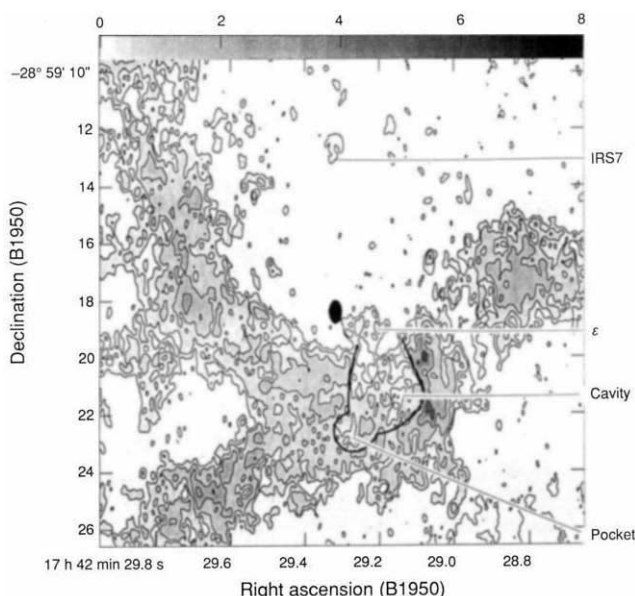


FIG. 2 a, A grey-scale contour map of the central $17''$ of the Galactic Centre at 2 cm with full width at half maximum $0.22'' \times 0.12''$ (position angle -0.22°). The final combined database has reasonably well sampled visibility (u-v) coverage with an equivalent total integration time of 3.5 h. Contours are $-2, 4, 8, 16, 32, 64, 256, 512 \times (8 \times 10^{-5})$ Jy per beam. The peak intensity is 1.6 Jy per beam. Grey-scale flux range is from 0 to 8 mJy per beam.

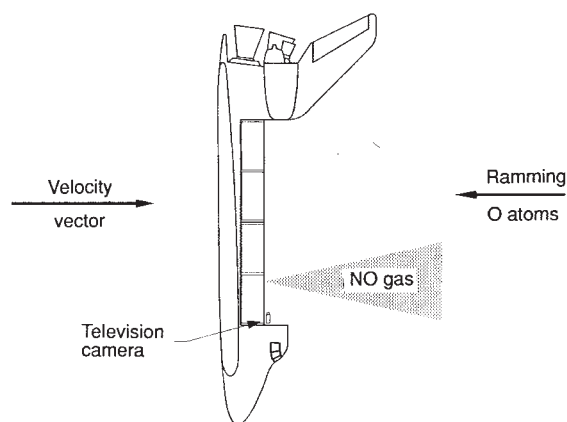


FIG. 1 Schematic of space shuttle *Discovery*, showing the velocity vector and the direction of release of gases from the gas canister used in the critical ionization velocity experiment.

on the shuttle's surface. This forms excited NO_2 , which radiates light as it desorbs^{6,7}. On a recent shuttle mission (STS-39) four gases, NO , CO_2 , Xe and Ne were released for a plasma experiment. Unintentionally, enough gas was scattered onto the surfaces of the shuttle tail that when NO was released a much more intense version of shuttle glow was observed. The other gases did not affect the normal shuttle glow. Under normal conditions the adsorbed NO that causes the glow probably comes either from the ambient atmosphere⁶ or from reactions in exhaust gases from the shuttle thrusters^{14,15}.

Spectroscopy of the shuttle glow¹¹ indicates that the spectrum is shifted to the red when compared with the continuum observed in the gas phase termolecular recombination of O and NO ; related laboratory studies of the reaction of hyperthermal O atoms with surface NO produced a variety of results. In one case a spectrum very similar to that of shuttle glow was seen^{16,17}, but in other experiments, spectra redder¹⁸ or bluer¹⁹ than the glow were observed. The activation energy for the glow, derived from a study¹² of the glow intensity of different space-borne materials, was estimated to be ~ 0.14 eV. Mass spectroscopic¹³ detection of NO_2 desorbed from shuttle surfaces facing the ram

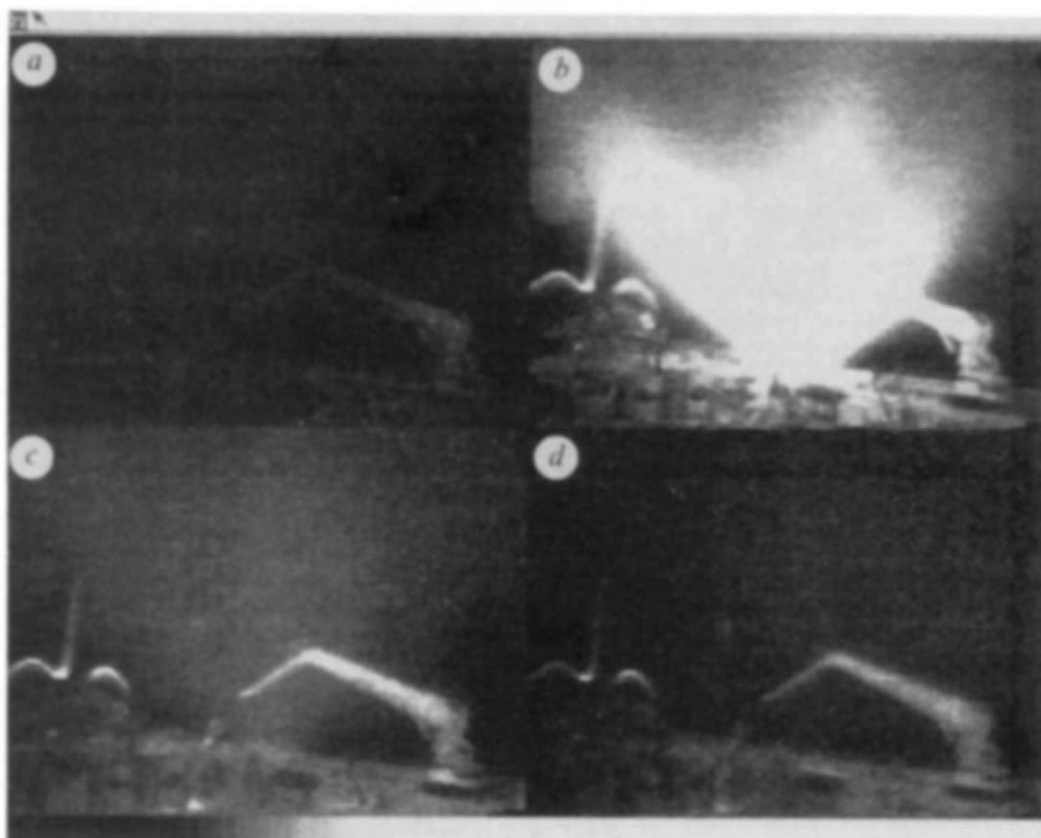
direction have confirmed the presence of sufficient NO_2 to account for the glow.

These pieces of evidence amounted to indirect support for the NO hypothesis, but experiments performed for the Strategic Defense Initiative Office on shuttle mission STS-39, whose main purpose was to investigate the critical ionization velocity theory for the different gases released, provided an opportunity for a more critical test. In the experiment^{20,21}, a mixture of dimers and monomers of the four gases was released. We do not discuss the critical ionization velocity experiment itself, but report visual observations of the intensity of the shuttle glow as influenced by the release of NO .

NO was released for 10 s from a high-pressure source through a nozzle which had 27 holes, each of diameter 0.1 mm. The holes were equally distributed on a plate. Figure 1 shows a sketch of the location of the gases in the shuttle bay, the velocity vector, and the direction of release.

Two effects were seen when NO was released: a bright blue-white plume formed in the gas phase, and the glow of the tail section of the space shuttle, as seen by the naked eye, was enhanced greatly. The intensity increased while the colour was

FIG. 2 Digitized picture from the in-bay TV cameras showing the effect of NO on the glow of the tail section and on the glow of the RMS. These images were taken before, during and after the NO release. *a*, The glow before the release; *b*, enhanced glow of the tail and RMS and the gaseous chemiluminescent reaction of NO ; *c*, the bay and tail 0.1 s after the NO release is turned off; *d*, the bay and tail 1 s after the release is turned off.



seemingly unaffected. The first is undoubtedly due to the reaction of $(\text{NO})_2$ with O:



This reaction has been postulated in a number of atmospheric NO releases from rocket-borne payloads^{22,23}, and has also been measured in laboratory experiments¹⁶. It will not be discussed further.

The released NO expands tens of metres before being collisionally back-scattered onto shuttle surfaces. The centre-of-mass energy in these collisions with atmospheric oxygen is insufficient for electronic excitation or dissociation of NO.

The second observation, which is our present subject, provides qualitative evidence that the visible portion of the shuttle glow is caused by recombination of hyperthermal O with surface-adsorbed NO, followed by relaxation of the excited NO_2^* :

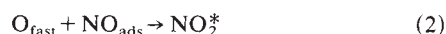


Figure 2 shows four video frames taken from the forward part of the shuttle bay with an intensified black and white video camera. The four panels show: (a), the quiescent glow on the ram side of the tail and of the remote manipulator arm (RMS) just before NO release; (b), the enhanced glow of the tail and the RMS, as well as a chemiluminescent gas-phase stream (the latter due to reaction 1), ~1 s into the release; (c), glow of the tail taken ~0.07 s after the NO flow is shut off; and (d), glow of the tail 1 s after the NO flow is turned off. In fact, Mende and Swenson²⁴ report that it takes almost 3 s before the enhanced tail glow disappears after thrusters are fired.

We believe this qualitative observation confirms the attribution of the visible shuttle glow to the sequence of reactions (2) and (3). Other surface-assisted recombination reactions may well occur to cause glow in the ultraviolet and infrared regions of the spectrum. If NO is responsible for the quiescent glow, as the observation implies, then this raises the question of the source of NO. A possible source is from thruster firings, which are used to control attitude and to perform manoeuvres. If that is the case, then the validity of thermodynamic equilibrium models of thruster exhausts must be questioned, as these models do not predict NO (see, for example, ref. 25). As mentioned earlier, NO can also be formed in the gas phase by the mechanism¹⁵



The NO will then migrate and adsorb on the surfaces. Subsequent O atom collisions produce NO_2^* emissions through reactions (2) and (3).

The glow decays exponentially after release with a characteristic time of 4 s to fall by a factor of e . The glow approaches a baseline (steady state) intensity asymptotically. Correction for this baseline yields a 3-s decay time. The observed decay represents the sum of atmospheric (O atom) scouring and physical desorption of the NO molecule from the shuttle tile surfaces. The observed decay is in accord with previous observations²⁴ and with the scouring time derived from laboratory measurements of fast atomic oxygen interacting with NO adsorbed on a variety of surfaces²⁶.

To confirm that the reaction of O with NO on surfaces is the source of the visible shuttle glow, work is needed to model the flux of NO onto the tail pod surface and to investigate the role of atomic oxygen velocity in the production of the visible glow. □

5. Garrett, H. B., Chutjian, A. & Gabriel, S. *J. Spacecraft Rockets* **25**, 321–341 (1988).
6. Swenson, G. R., Mende, S. B. & Clifton, K. S. *Geophys. Res. Lett.* **12**, 97–100 (1985).
7. Heppner, J. P. & Meredith, L. H. *J. Geophys. Res.* **12**, 51–65 (1958).
8. Torr, M. R., Hays, P. B., Kennedy, B. C. & Walker, J. C. C. *Planet. Space Sci.* **25**, 173–184 (1977).
9. Yee, J. H. & Abreu, V. J. *Proc. SPIE Tech. Symp.* **338** (ed. Maag, C. R.) 120 (SPIE, Bellingham, 1982).
10. Yee, J. H., Abreu, V. J. & Daigarno, A. *Geophys. Res. Lett.* **11**, 1192–1194 (1984).
11. Paulsen, D. E., Sheridan, W. F. & Huffman, R. E. *J. chem. Phys.* **53**, 647–658 (1970).
12. Swenson, G. R., Mende, S. B. & Meyerott, R. E. *Nature* **323**, 519–522 (1986).
13. von Zahn, U. & Murad, E. *Nature* **321**, 147–148 (1986).
14. Meyerott, R. E. & Swenson, G. R. *Planet. Space Sci.* **38**, 555–566 (1990).
15. Green, B. D., Rawlins, W. T. & Marinelli, W. J. *Planet. Space Sci.* **34**, 879–887 (1986).
16. Caledonia, G. E. *et al. Geophys. Res. Lett.* **17**, 1881–1884 (1990).
17. Swenson, G. R., Leone, A., Holtzclaw, K. W. & Caledonia, G. E. *J. geophys. Res.* **96**, 7603–7612 (1991).
18. Arnold, G. S. & Coleman, D. J. *J. chem. Phys.* **88**, 7147–7156 (1988).
19. Orient, O. J., Chutjian, A. & Murad, E. *Phys. Rev. A* **41**, 4106–4108 (1990).
20. Murad, E., Lai, S. T. & Stair, A. T. Jr *J. geophys. Res.* **91**, 10188–10192 (1986).
21. Green, B. D. *et al. 26th Aerospace Sciences Meeting Paper AIAA-88-0432* (AIAA, Washington DC, 1988).
22. Fontijn, A. & Rosner, D. E. *J. chem. Phys.* **46**, 3275–3276 (1967).
23. Golomb, D. & Good, R. E. *J. chem. Phys.* **49**, 4176–4180 (1968).
24. Mende, S. B., Swenson, G. R. & Llewellyn, E. J. *Adv. Space Sci.* **7**, 169–178 (1987).
25. Pickett, J. S., Murphy, G. R., Kurth, W. S., Goertz, C. K. & Shawhan, S. D. *J. geophys. Res.* **A90**, 3487–3497 (1985).
26. Caledonia, G. E., Holtzclaw, K. W., Sonenfroth, D. & Krech, R. in *Proc. Workshop on Induced Environment of Space Station Freedom* (ed. Torr, M. R.) (NASA, Huntsville, 1990).

ACKNOWLEDGEMENTS. We thank SDIO and M. Harrison, manager of the IBSS program, for supporting the gas release experiment on STS-39.

Linking of vortex rings

Hassan Aref & Ireneusz Zawadzki

University of California at San Diego, La Jolla, California 92093-0225, USA

THE topology of vortex lines, which trace the local vorticity in a fluid flow just as streamlines trace the velocity, is important in attempts to understand, describe and control flows in various applications. Changes in this topology may, for example, affect mixing in flows, and may be significant for the dynamics of turbulence. In particular, the behaviour of closed loops of vortex lines (vortex rings) has been studied ever since Kelvin's 'vortex atom' theory, and contributed to Tait's development of the topological theory of knots. Several recent experimental and computational studies^{1–9} have explored the 'reconnection' of initially distinct vortex rings; particularly elegant are Schatzle's¹⁰ experiments in which two vortex rings, inclined towards one another, go through two reconnections, after which two new rings, comprising half of each of the originals, emerge. Here we describe three-dimensional numerical simulations which establish a simple mechanism by which the linking of two vortex rings may be achieved starting from an unlinked initial state. Appropriate initial states were identified by simulating the unlinking of two initially linked vortex rings and then reversing the vorticity of the final state and running the simulation backwards. This computational procedure sheds light on why, both in experiment and simulation, linking is not always achieved from an arbitrary initial configuration of unlinked vortex rings set on a collision course.

A measure of the linking of vortex lines is provided by the helicity integral¹¹

$$H = \int \mathbf{u} \cdot \boldsymbol{\omega} \, dV \quad (1)$$

where $\mathbf{u}$ is the velocity field, $\boldsymbol{\omega} = \nabla \times \mathbf{u}$ the vorticity, and the integral is extended over the entire flow volume. For vortex filaments concentrated on space curves, equation (1) reduces to the topological linkage integral or 'winding number' studied already by Gauss. In inviscid flow the helicity is an integral of the motion. Thus, vortices that are unlinked initially must remain so. In viscous flow, however, 'vortex reconnections' can occur, leading to changes in the topology of the vortex lines and to fluctuations in the value of the helicity.

We have conducted numerical simulations of this type of flow using a three-dimensional eulerian-lagrangian, vortex-in-cell (VIC) code. (Refs 12 and 13 give a discussion of three-dimensional 'vortex methods'.) VIC leads to a nearly inviscid

Received 2 July; accepted 18 September 1991.

1. Banks, P. M., Williamson, P. R. & Raitt, W. J. *Geophys. Res. Lett.* **10**, 118–121 (1983).
2. Mende, S. B., Garriott, O. K. & Banks, P. M. *Geophys. Res. Lett.* **10**, 122–125 (1983).
3. Green, B. D., Caledonia, G. E. & Wilkerson, T. D. *J. Spacecraft Rockets* **22**, 500–511 (1985).
4. Murad, E., in *Physics of Space Plasmas (1985–1987)*, Vol. 6 (eds Chang, T. *et al.*) 147–161 (Scientific, Cambridge, MA, 1987).

FIG. 1 Simulation of the collision of identical vortex rings. Perspective view of isosurfaces of vorticity magnitude at nine instants complementing the experimental pictures by Schatzle¹⁰ (time progresses from *a* to *i*). Note the near inversion symmetry of the panels about the central one. (Isosurfaces of vorticity magnitude are at 20% of the maximum.)

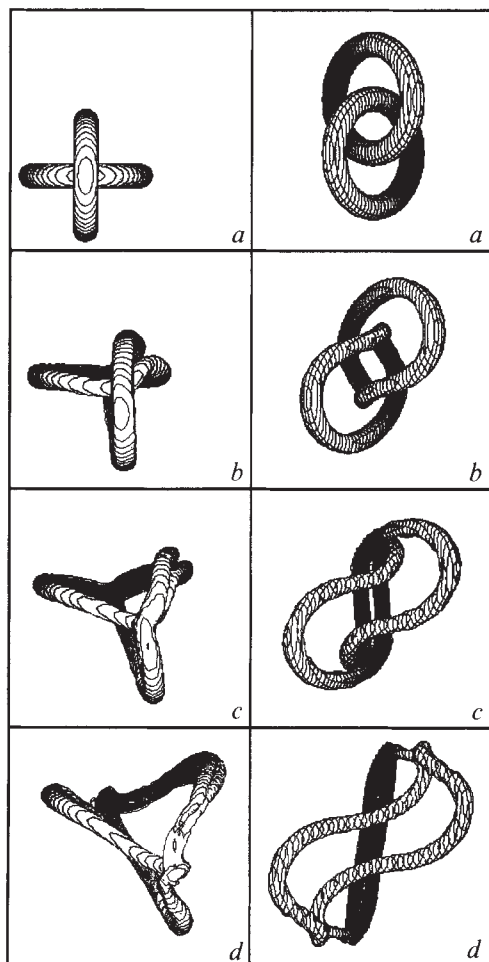
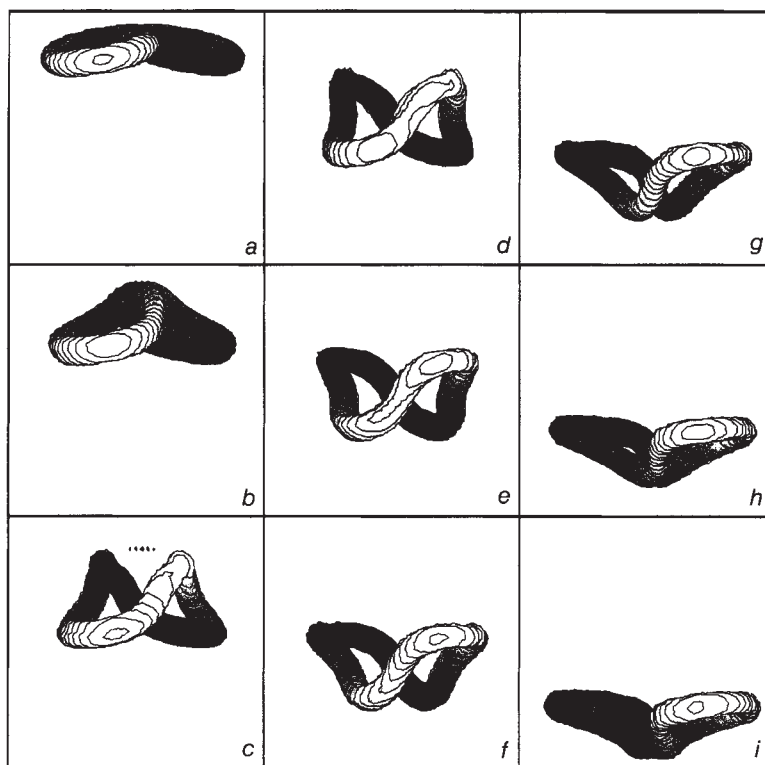


FIG. 2 Two views of the unlinking of initially linked vortex rings situated in perpendicular planes. The end result, which would occur somewhat after the final panel, is one large ring. Considerable annihilation of close, opposite-signed vorticity takes place in the double-strand seen in *d*. This evolution is not a good candidate for reversal to produce linking. (Isosurfaces of vorticity magnitude at 20% of the maximum.)

flow representation, but because of vorticity diffusion on the underlying grid, vortex reconnection can take place. Extensive calculations with this code¹⁴ suggest that it is extremely useful for following interactions of spatially localized vortices, such as vortex rings and tubes. With a 64^3 grid and 25,000 vortex particles, simulations of flows at Reynolds number of the order of 1,000 or so are possible. In particular, we can simulate the reconnection processes observed by Schatzle¹⁰ with considerable fidelity (Fig. 1).

Stimulated by this success in reproducing known patterns of motion, we have explored the possibility of making two vortex rings link as they evolve in time. This would provide a specific mechanistic model of the production of helicity, although, as Moffatt has already stressed¹¹, and as the numerical experiments bear out, much of the important helicity dynamics involves rearrangement of distributed vorticity, for example, the twisting of vortex lines within individual vortices. Several attempts at colliding circular vortex rings with one another to produce a linked configuration failed to yield the desired result. The rings would deform considerably, but they refused to link.

While conducting simulations of Schatzle's experiment¹⁰, we noticed that the entire process has considerable 'reversibility'. Thus, if we take the final state, Fig. 1*i*, with two emerging rings, reverse all the vorticity and run it backwards, we reproduce a state very close to the initial one, Fig. 1*a*. In fact, a sequence of snapshots of the motion ordered backwards in time looks very similar to the sequence of snapshots of the forward evolution rotated 180° . (The reader is invited to rotate Fig. 1 by 180° and compare the result with the original figure.) We therefore decided to start with a configuration in which we had two linked rings, run it forward until it produced two unlinked rings and then use that final state with its vorticity reversed as the initial state in a linking simulation.

This strategy showed that for many initial conditions the unlinking of rings proceeds through states that involve considerable annihilation of opposite-signed vorticity. Such processes clearly cannot be reversed in time, as we would then be demanding the creation of tightly bound, opposite vortices in the interior of the flow. Hence, these particular initial conditions, which

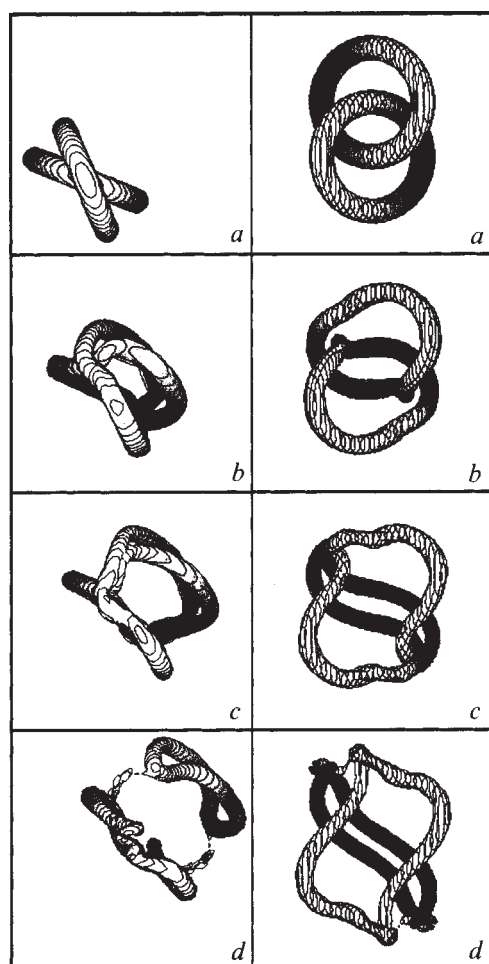


FIG. 3 Two views of the unlinking of two initially linked vortex rings situated in planes making an angle of 45° . The end result is two elongated rings situated, approximately, in parallel planes. Two simultaneous, localized reconnections occur. This evolution is a good candidate for reversal to produce linking. (Isosurfaces of vorticity magnitude at 20% of the maximum.)

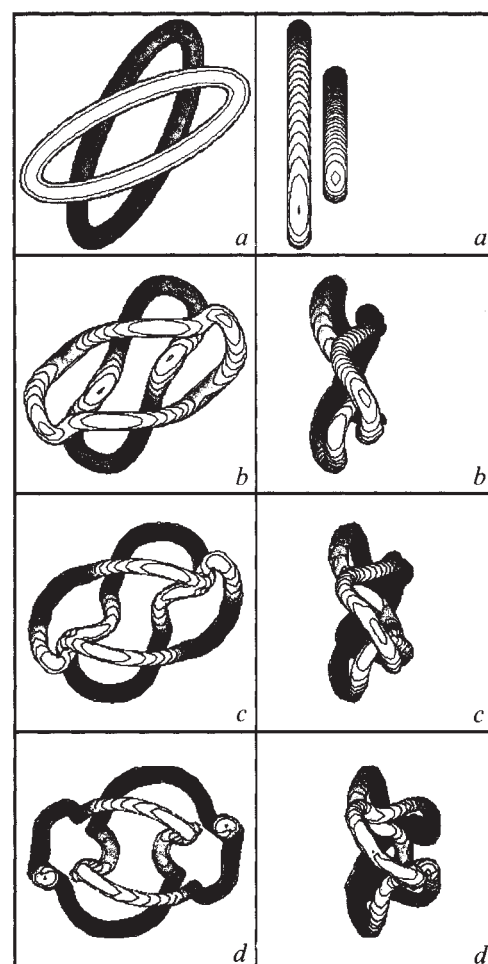


FIG. 4 Perpendicular views of the linking of two unlinked, elliptical vortex rings situated initially in parallel planes. The initial state is an idealization of the final state in Fig. 3 (with vorticity reversed). The helicity grows from zero to 65% of the 'linkage integral' value. (Isosurfaces of vorticity magnitude at 20% of the maximum.)

include two identical rings in perpendicular planes, linked so that each passes through the centre of the other (Fig. 2), are not useful for generating unlinked configurations that will lead to linking. In Fig. 2, the stretched double-strand seen in the 'background' in Fig. 2d (right column) will annihilate because of viscous effects, and a single, elongated ring will be left. The alignment of vorticity of opposite sign was observed by Siggia¹⁵ in model calculations of the evolution of a single vortex loop. Although this effect also occurs in two-ring interactions such as the one shown in Fig. 2, our objective was to find general classes of motion in which it is avoided.

After some experimentation with initial states of the type shown in Fig. 2, we found cases, such as the one shown in Fig. 3, where the initial state is linked, and the end state consists of two separate ring vortices. This is a candidate for reversal. The simplest initial condition that resembles the final geometry in Fig. 3 and produces linking consists of two elliptical vortex rings in parallel planes with major axes inclined by $\sim 45^\circ$ (Fig. 4a; the actual angle of inclination is 40°). This state may be viewed as an idealized version of the state in Fig. 3d. The orientation of the vorticity in the rings should be such that they move in the same direction. If the major axes are parallel as the vortices approach each other closely, a 'leapfrogging' akin to what is observed for coaxial, circular, vortex rings¹⁶ can take place. If the major axes are perpendicular, the elliptical vortices reconnect at four points and two coaxial rings are produced. The

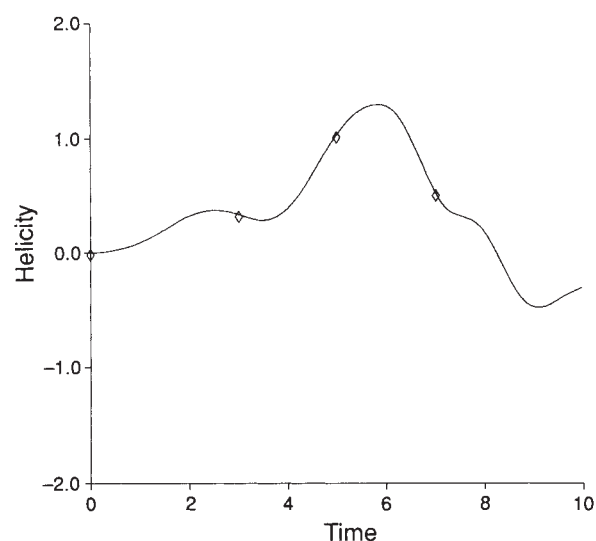


FIG. 5 Graph of the helicity as a function of time for the simulation in Fig. 4. Instants corresponding to the four panels are indicated by symbols. The maximum helicity, which occurs for a configuration intermediate between Fig. 4c and d, is 65% of the full linkage value. (Units of length and time are chosen such that the computational domain has side 2π and the circulation of each ring initially is 1.)

essence of the initial condition in Fig. 4a is that the elliptical rings first collide and reconnect at just two points. Clearly, this can occur for a broad range of initial configurations; for example, the angle need not be exactly 40°, and the two elliptical rings need not be identical. The initial spacing of the planes of the two rings is important, however.

Figure 4 traces the evolution. It is clear that in the state in Fig. 4d the two new rings formed, each of which consists of half of each of the original vortices, are linked. This can only occur for a finite time. By the reversibility argument we have been using to construct the initial condition, the linked rings will unlink again. We have calculated the helicity continuously during the interaction in Fig. 4, and it is plotted in Fig. 5. The symbols designate states corresponding to the panels in Fig. 4. The maximum helicity, which occurs somewhere between Fig. 4c and d, is 65% of the linkage integral value. (The helicity of the initial state, Fig. 4a, vanishes.) In similar simulations we have seen maxima of helicity of 70% of the linkage value. Even though Fig. 4d indicates that the full linkage value might be recovered, a considerable amount of helicity is associated with the winding of vortex lines within the cores of individual rings, and with vortex lines that thread both vortex rings even after the reconnection. Only in the limit of concentrated line vortices would all the helicity be associated with the topological linkage.

Thus, we have identified a class of initial conditions that lead to linking of vortices in nearly inviscid flow. The mechanism, of interest in its own right, is potentially of importance in other flow situations involving concentrated vortices. □

Received 9 May; accepted 3 September 1991.

1. Kambe, T. & Takao, T. *J. phys. Soc. Japan* **31**, 591–599 (1971).
2. Oshima, Y. & Asaka, S. *Nat. Sci. Rep. Ochanomizu Univ.* **26**, 31–37 (1975); *J. phys. Soc. Japan* **42**, 708–713 (1977).
3. Ashurst, W. & Meiron, D. I. *Phys. Rev. Lett.* **58**, 1632–1635 (1987).
4. Melander, M. & Zabusky, N. J. *Fluid Dyn. Res.* **3**, 247–250 (1988).
5. Kida, S. & Takaoka, M. *Fluid Dyn. Res.* **3**, 257–261 (1988).
6. Oshima, Y. & Izutsu, N. *Phys. Fluids* **31**, 2401–2403 (1988).
7. Melander, M. V. & Hussain, F. in *Topological Fluid Mechanics* (eds Moffatt, H. K. & Tsinober, A.) 485–499 (Cambridge University Press, 1990).
8. Kida, S., Takaoka, M. & Hussain, F. *Phys. Fluids A1*, 630–633 (1989); (Corrigendum) *ibid.* **2**, 638 (1990); *J. Fluid Mech.* **230**, 583–646 (1991).
9. Melander, M. V. & Hussain, F. *Phys. Fluids A1*, 633–636 (1989).
10. Schatzle, P. R. thesis, California Institute of Technology (1987).
11. Moffatt, H. K. *J. Fluid Mech.* **35**, 117–129 (1969).
12. Leonard, A. *Ann. Rev. Fluid Mech.* **17**, 523–559 (1985).
13. Winckelmans, G. & Leonard, A. in *Mathematical Aspects of Vortex Dynamics* Ch. 2 (ed. Caflisch, R. E.) (Society of Industrial and Applied Mathematics, 1989).
14. Zawadzki, I. & Aref, H. *Phys. Fluids A3*, 1405–1410 (1991); *J. comput. Phys.* (submitted).
15. Siggia, E. D. *Phys. Fluids* **28**, 794–805 (1985).
16. Yamada, H. & Matsui, T. *Phys. Fluids* **21**, 292–294 (1978).

ACKNOWLEDGEMENTS. This paper is dedicated to the memory of P. R. Schatzle, whose experiments stimulated our interest in vortex ring interactions. This research was supported by the DARPA University Research Initiative programme. Computations and visualizations were performed at the San Diego Supercomputer Center, which is supported by the NSF.

Microwave localization by two-dimensional random scattering

Rachida Dalichaouch*, J. P. Armstrong*, S. Schultz*, P. M. Platzman† & S. L. McCall†

* University of California, San Diego, Physics 0319, 9500 Gilman Drive, La Jolla, California 92093-0319, USA

† AT&T Bell Laboratories, 600 Mountain Avenue, Murray Hill, New Jersey 07974, USA

WAVEFUNCTIONS of electrons or photons in a strongly scattering random medium may become localized owing to the underlying wave nature of the particles^{1,2}. Particularly surprising and counter-intuitive is the prediction that, under appropriate conditions, scatterers placed randomly in space will always produce fully localized states—that is, an energy distribution of the normal modes whose envelope decays exponentially in all directions. In consequence, energy at the resonant frequency of a localized mode, injected into that mode's region of space, cannot diffuse away, but remains trapped until dissipated. Here we report measurements of the electric-field energy density for microwave radiation localized in essentially two-dimensional space by scattering from a random array of dielectric cylinders placed between a pair of parallel conducting plates. We detect regions of high energy density representing the signature of localized modes. The available range of measured variables, scattering materials and cylinder configurations offer the opportunity to provide quantitative answers to important general questions about strong localization. In particular, a better understanding of two-dimensional localization raises the possibility of using localized-mode resonances as a diagnostic tool for situations in which localization phenomenon may occur naturally³—for example, in investigations of the internal distribution of media and defects in geological strata, under-ocean topology or electronic thin films, all of which may exhibit pseudo-two-dimensional characteristics.

For a simple physical explanation of the origins of the localization, and an illustration of its effects on transport properties, we consider resistivity. We may first see the conduction electrons in a suitable amorphous metal alloy (at cryogenic temperatures)

as undergoing a random walk as they scatter elastically (without energy loss) from atom to atom. One statistical description of this random walk is by the corresponding diffusion constant. If the particles were classical, one could predict how far a given group of electrons would diffuse in a given time. It is highly likely that a given particle during its classical random walk will return to its starting point. In one dimension (1D) and two dimensions (2D) this is always so, and sometimes so, in three dimensions (3D), and we must take into account that the particle could arrive at that return point by two corresponding time-reversed scattering paths. As the underlying description is based on a wave equation, there is then the possibility of interference. The two wave amplitudes at the return point in space and time must be added and then squared to get the total probability density. If the waves add constructively, the probability for the particle to have returned to that point will be greater than would have been predicted classically, with a consequent reduction in the diffusion constant. Conditions in which such interference effects have small, but easily measured, consequences for particle transport⁴ or optical coherent backscattering^{5,6} are called weak localization. Striking effects, such as the particle self-interfering so strongly that it never diffuses away, but is 'localized' in space, can also occur, and therefore the diffusion constant approaches zero. This condition is termed strong localization (SL), and the characteristic length over which the particle probability density decays exponentially in space is called the localization length.

In some circumstances, as the energy of the particle is changed, the diffusion constant can change rapidly from near zero (localized) to finite (delocalized conditions). Resistance measurements would indicate that the material underwent a transition from insulator to metal; the energy at which the transition occurs is termed a mobility edge². In 3D, such mobility edges are known to exist¹⁸. In contrast, in 1D and 2D it is generally believed that in any sufficiently large system, with any amount of randomness, all states are localized. Some controversy still exists, however, and other interesting questions remain to be tested. Model experimental SL systems, where parameters can be readily changed, and for which exact corresponding calculations can be performed, without, for example, the problems associated with particle-particle interactions, would be a valuable tool.

Although the bulk of the effort toward understanding SL has

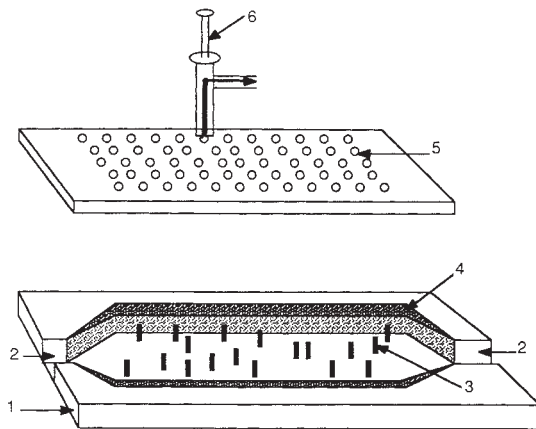


FIG. 1 Schematic representation of the scattering system (not to scale). (1) Bottom and sides of scattering chamber (made as one piece), (2) inlet and outlet waveguide ports, (3) ~500 dielectric cylinders randomly placed on a square lattice, (4) microwave absorber, (5) coverplate with holes for sampling electric energy density and (6) a tuned microwave probe which is placed in the sampling holes in the coverplate, and whose output is measured by a homodyne detector at two phases 90° apart.

been oriented towards quantum mechanical systems^{4,7}, the basic effect is a consequence of wave interference, and can therefore be described within the framework of classical wave phenomena^{3,5,6,8-10}. Possible applications and detailed consequences for strongly localized photons will, of course, be different from those of electrons^{11,12}. We have made a model 2D electromagnetic system whose properties, in principle, are completely described by solutions of Maxwell's equations for the electromagnetic fields. Related microwave localization experiments in 3D have used random dielectric and metal spheres¹³, and a specific periodic dielectric structure¹⁴.

To establish well-defined localized states with reasonable localization lengths, we require a random set of scatterers in a system with the following properties: (1) the inelastic mean free path, l_i , should greatly exceed the elastic (kinetic-energy conserving) mean free path, l_e , for the modes to be well resolved. (2) l_e should be small, and if possible comparable to the distance between scatterers (this condition requires a large dielectric mismatch). (3) The Ioffe-Regle (strong localization) condition, $2\pi l_e/\lambda = kl_e \approx 1$, should be satisfied, where λ is the wavelength at the frequency of interest. We also require (4) a procedure and apparatus for identifying the excitation frequency of the localized modes; (5) a means of establishing the location of a frequency-selected localized mode; (6) a means of measuring the wavefunction with a desired spatial resolution; and (7) criteria and corresponding tests to verify that the measured 'mode' is a localized mode, and not an artefact.

In our strongly scattering system, we can produce, identify and select localized modes caused by random scatterers, and then obtain a complete spatial mapping of the complex electric field—that is, the photon wavefunction as a function of all system parameters. We create localized states of microwave energy in a set of dielectric cylinders placed between two parallel aluminium plates which set up a 2D scattering chamber (Fig. 1). The cylinders are placed randomly on half the sites of a square lattice whose unit cell has a lattice constant of $a_0 = 1.27$ cm.

We use three configurations of excitation and detection of the microwave fields for the process of identification, selection and mapping. In configuration 1, swept frequency-modulated microwave power is injected through one tapered port and detected at another, thereby measuring the frequency-dependent transmissivity of the array. Configuration 2 is similar to 1 except that the microwave power is injected through one of a number of small holes in either of the aluminium plates. In configuration 3, the microwave frequency is constant, and power may be injected into the chamber as in 1 or 2, but the fields are detected by a tuned probe placed in one of a set of 2-mm sampling holes drilled into the chamber coverplate and spaced 1.27 cm apart on a square lattice. The coverplate can be mechanically translated more than 1.27 cm so one can scan the entire chamber surface. The amplitude and phase of the microwave field emanating from the probe are determined by homodyne detection using a reference whose phase may be shifted by 90° . In a 1D version of this work¹⁵, which incorporates random dielectric slabs in a

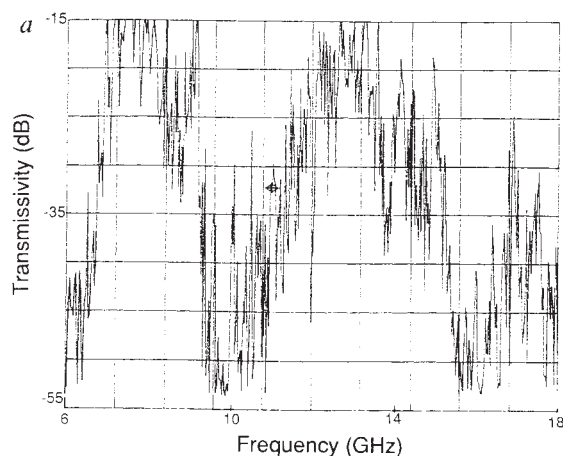
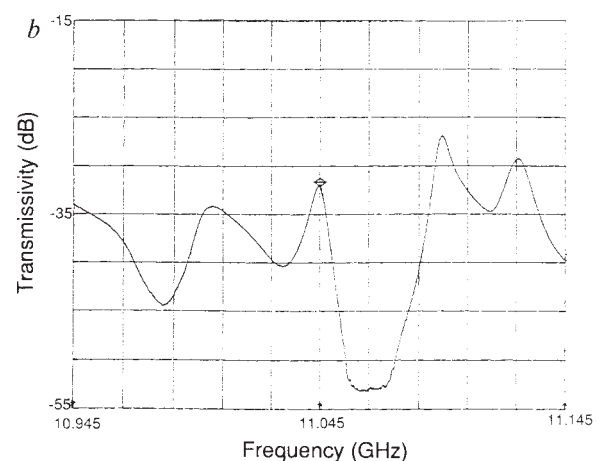


FIG. 2 The ratio of transmitted to incident power as a function of frequency for a set of dielectric cylinders ($\epsilon = 9$), of radius 0.98 cm, placed randomly on half the sites of a square lattice of spacing $a_0 = 1.27$ cm. *a*, An extended sweep range of 6–18 GHz. The sharp peaks are attributed to excitation of localized modes. The superimposed modulation with width of ~4 GHz is



attributed to vestiges of the stop and pass bands of the underlying lattice^{15,16}. *b*, A section of the data in *a*, expanded to display a peak centred at 11.045 GHz. Damping tests described in the text identified this peak as being associated with a localized mode centrally located in the chamber.

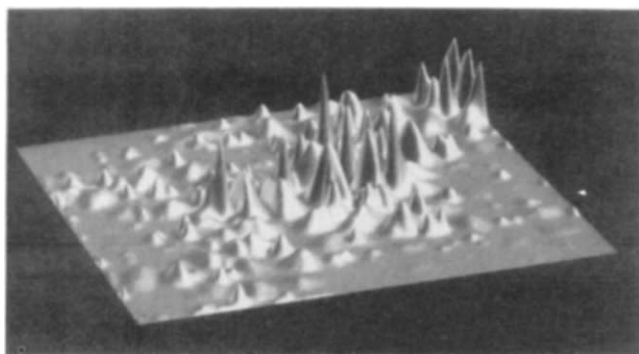


FIG. 3 A perspective plot of the absolute square of the complex field emanating from the microwave probe (proportional to the electric energy density, ϵE^2) as a function of space for a $14a_0 \times 14a_0$ region located near the centre of the $36a_0 \times 27a_0$ scattering array. The most frequent peak width is roughly half a wavelength, and simply represents the standing wave interferences. The cluster of peaks in the centre of the figure all belong to the same localized mode. The overall radial amplitude dependence from this centre cluster sets the characteristic scale for the localization length. The cluster of six peaks in the upper right-hand corner is part of the tail of a neighbouring localized mode. The tests performed to verify that the central cluster of peaks was due to a single localized mode are described in the text.

single-mode waveguide, we have determined that the power detected at the output of the probe for configuration 3 is proportional to ϵE^2 , where ϵ is the dielectric constant and E the electric field inside the slab material under the probe tip. This calibration was established by comparing the measured spatial dependence with numerical simulations, which can be accurately calculated in 1D.

By inserting a capillary filled with water into one or more of the sampling holes, we can locally damp the electric field, selectively attenuating spatially separated modes. This allows us to find the modes, and then apply a series of tests.

Figure 2a shows the transmissivity obtained using configuration 1. We note the dense sharp peak structure with a superimposed modulation which we identify as due to vestiges of the stop and pass bands of the underlying lattice^{15,16}. (The stop and pass bands, which are the consequences of wave interference, are those regions of frequency where a wave incident on a perfect lattice is, respectively, strongly reflected and transmitted. An analogous example is that X-rays of energies that satisfy the Bragg scattering condition occur within a stop band.) Figure 2b shows part of the data of Fig. 2a at much higher resolution. The observed frequency width of the peaks is consistent with the known losses of the aluminium and dielectric material. The transmission peaks in Fig. 2b are not well resolved, suggesting that there are several localized modes excited in the sample at any given frequency. For the size of the chamber, a simple calculation of the 2D density of states for a box filled with the average ϵ implies that we would expect three or four modes within the frequency width of the peaks in the transmission data. In these 2D experiments at microwave frequencies, because we can readily satisfy the strong scattering conditions $kl_e \approx ka_0 \approx 1$, we can achieve localization without needing to work near the band edges of an underlying lattice¹⁷. For example, we have examined the peak structure when the dielectric cylinders are laid on their sides and then randomly distributed and oriented, without any underlying lattice, and observed localized modes of similar extent.

To study one of the modes least affected by boundaries, that is, one localized near the centre of the array, we use a simple selection procedure. We ascertain the qualitative effect on the transmission spectrum of a water filled capillary inserted in the central region. The peak at 11.045 GHz in Fig. 2b was severely attenuated, whereas neighbouring ones were not. We then lock the source frequency to the centre of the selected resonant peak and complete a spatial mapping using configuration 3.

In the type 3 configuration we measure both components of the complex field coupled into the probe. We form the absolute square of these, which, as noted, is proportional to the electric energy density ϵE^2 in the dielectric at the probed position. Figure 3 shows a perspective plot of ϵE^2 as a function of position for an excitation frequency of 11.045 GHz. The region shown represents a section of only $14a_0 \times 14a_0$, near the centre of the total array of $36a_0 \times 27a_0$. There are oscillations on the scale of half the wavelength because of the standing wave character of the mode. As the microwave magnetic field is zero when the

curl of E is zero, the maxima of the peaks in Fig. 3 are points of total energy density. Thus, the surface enveloping all such points is a fairly good representation of the falloff of the total energy density in all directions. As E is continuous, there are additional sharp changes in ϵE^2 at the cylinder boundaries. The overall amplitude decay for the cluster of peaks in the centre of the figure sets a characteristic scale for the localization length.

Also shown in Fig. 3 is the tail of another mode, excited slightly off its resonant frequency. An important aspect of this work is the testing procedure for mode discrimination and definition, which we now describe. (1) Frequency dependence. In a slight modification of configuration 3, we monitor the output of the probe while it is stationed over a peak in ϵE^2 anywhere within the localized mode, while slightly detuning the frequency. We find all peaks associated with a given mode are maximized at the same frequency. Similarly, the peaks attributed to a neighbouring mode tune together, but are maximal only at that mode's resonant frequency. (2) Damping studies. Whenever a water-filled capillary is placed in the region of a single peak within a given localized mode, all the other peaks in that mode are also strongly damped. In contrast, peaks of other localized modes, even those close by, are only slightly damped. (3) Direct mode excitation. In a variation of configuration 2, microwave power at the resonant frequency is injected into any strong peak of a given localized mode, and we find that all peaks in that mode are suitably excited, whereas nearby modes are minimally excited. (4) Independence of boundary conditions. Most striking is that the peak structure in the region identified as a localized mode in Fig. 3, is essentially independent of the boundary conditions. For example, the data of Fig. 3 were obtained for a lattice of $36a_0 \times 27a_0$. When cylinders are removed from the exterior regions to produce a square lattice of only $21a_0$, the localized mode amplitudes increased, but their pattern remained basically unchanged. □

Received 28 May; accepted 17 September 1991.

- Anderson, P. W. *Phys. Rev.* **109**, 1492-1505 (1958).
- Lee, P. A. & Ramakrishnan, T. V. *Rev. Mod. Phys.* **57**, 287-337 (1985).
- Ishimaru, A., *Wave Propagation and Scattering in Random Media* (Academic, New York 1978).
- Localization, Interaction, and Transport Phenomena (eds Kromer, B., Bergmann, G. & Broynsraede, Y.) (Springer, Berlin, 1985).
- van Albada, M. P. & Lagendijk, A. *Phys. Rev. Lett.* **55**, 2692-2695 (1985).
- Wolf, P. E. & Maret, G. *Phys. Rev. Lett.* **55**, 2696-2699 (1985).
- Abrahams, E. P., Anderson, W., Licciardello, D. C. & Ramakrishnan, T. V. *Phys. Rev. Lett.* **42**, 673-676 (1979).
- Directions in Condensed Matter Physics. Vol. 8: Scattering and Localization of Classical Waves in Random Medium* (ed. Sheng, P.) (World Scientific, Singapore, 1990).
- John, S. *Phys. Rev. Lett.* **53**, 2169-2172 (1984).
- Anderson, P. W. *Phil. Mag.* **B52**, 505-509 (1985).
- John, S. *Physics Today* **44**, 32-40 (1991).
- Yablonovitch, E. *Phys. Rev. Lett.* **58**, 2059-2062 (1987).
- Garcia, N. & Genack, A. *Phys. Rev. Lett.* **63**, 1678-1681 (1989).
- Yablonovitch, E. & Leung, K. M. *Nature* **351**, 278-278 (1991).
- Dalichaouch, R. thesis, Univ. of California at San Diego (1990).
- McCall, S. L., Platzman, P. M., Dalichaouch, R., Smith, D. R. & Schultz, S. *Phys. Rev. Lett.* (in the press).
- John, S. *Phys. Rev. Lett.* **58**, 2486-2490 (1987).
- Mott, N. F. *Adv. Phys.* **16**, 49-144 (1967).

ACKNOWLEDGEMENTS. We thank D. R. Smith and F. Mickadeit for their aid in the acquisition of the spatial mapping data, and D. Eigler for performing the image processing of the data shown in Fig. 3. The research at UCSD was supported by the NSF, Division of Materials Research.

Helical microtubules of graphitic carbon

Sumio Iijima

NEC Corporation, Fundamental Research Laboratories,
34 Miyukigaoka, Tsukuba, Ibaraki 305, Japan

THE synthesis of molecular carbon structures in the form of C_{60} and other fullerenes¹ has stimulated intense interest in the structures accessible to graphitic carbon sheets. Here I report the preparation of a new type of finite carbon structure consisting of needle-like tubes. Produced using an arc-discharge evaporation method similar to that used for fullerene synthesis, the needles grow at the negative end of the electrode used for the arc discharge. Electron microscopy reveals that each needle comprises coaxial tubes of graphitic sheets, ranging in number from 2 up to about 50. On each tube the carbon-atom hexagons are arranged in a helical fashion about the needle axis. The helical pitch varies from needle to needle and from tube to tube within a single needle. It appears that this helical structure may aid the growth process. The formation of these needles, ranging from a few to a few tens of nanometres in diameter, suggests that engineering of carbon structures should be possible on scales considerably greater than those relevant to the fullerenes.

Solids of elemental carbon in the sp^2 bonding state can form a variety of graphitic structures. Graphite filaments can be produced, for instance, when amorphous carbon filaments formed by thermal decomposition of hydrocarbon species are subsequently graphitized by heat treatment^{2,3}. Graphite filaments can also grow directly from the vapour-phase deposition of carbon^{4,5}, which also produces soot and other novel structures such as the C_{60} molecule⁶⁻⁸.

Graphitic carbon needles, ranging from 4 to 30 nm in diameter and up to 1 μm in length, were grown on the negative end of the carbon electrode used in the d.c. arc-discharge evaporation of carbon in an argon-filled vessel (100 torr). The gas pressure was much lower than that reported for the production of thicker

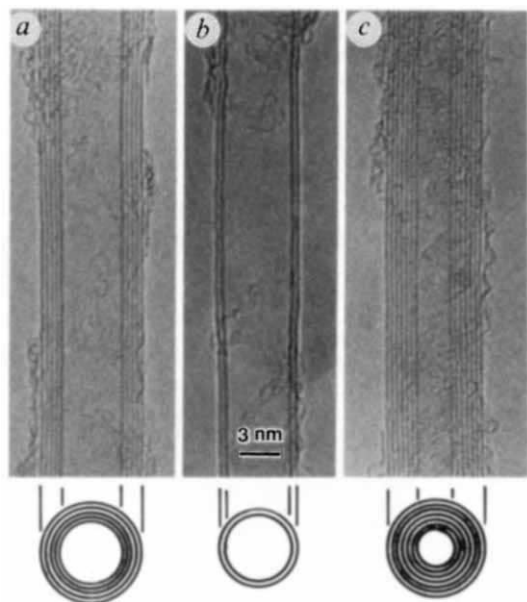


FIG. 1 Electron micrographs of microtubules of graphitic carbon. Parallel dark lines correspond to the (002) lattice images of graphite. A cross-section of each tubule is illustrated. *a*, Tube consisting of five graphitic sheets, diameter 6.7 nm. *b*, Two-sheet tube, diameter 5.5 nm. *c*, Seven-sheet tube, diameter 6.5 nm, which has the smallest hollow diameter (2.2 nm).

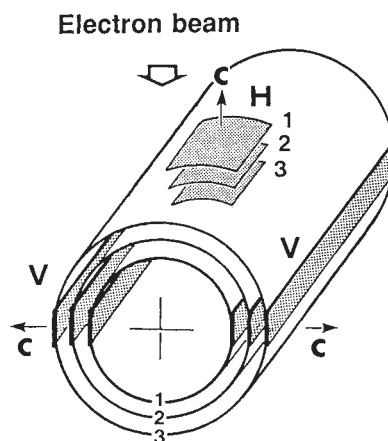


FIG. 2 Clinographic view of a possible structural model for a graphitic tubule. Each cylinder represents a coaxial closed layer of carbon hexagons. The meaning of the labels V and H is explained in the text.

graphite filaments⁵. The apparatus is very similar to that used for mass production of C_{60} (ref. 9). The needles seem to grow plentifully on only certain regions of the electrode. The electrode on which carbon was deposited also contained polyhedral particles with spherical shell structures, which were 5–20 nm in diameter. The needle structures were examined by transmission electron microscopy (electron energies of 200 keV).

High-resolution electron micrographs of typical needles show {002} lattice images of the graphite structure along the needle axes (Fig. 1). The appearance of the same number of lattice fringes from both sides of a needle suggests that it has a seamless and tubular structure. The thinnest needle, consisting of only two carbon-hexagon sheets (Fig. 1*b*), has an outer and inner tube, separated by a distance of 0.34 nm, which are 5.5 nm and 4.8 nm in diameter. The separation matches that in bulk graphite. Wall thicknesses of the tubules range from 2 to 50 sheets, but thicker tubules tend to be polygonized. This low dimensionality and cylindrical structure are extremely uncommon features in inorganic crystals, although cylindrical crystals such as serpentine¹⁰ do exist naturally.

The smallest tube observed was 2.2 nm in diameter and was the innermost tube in one of the needles (Fig. 1*c*). The diameter corresponds roughly to a ring of 30 carbon hexagons; this small diameter imposes strain on the planar bonds of the hexagons and this causes two neighbouring hexagons on the ring to meet at an angle of $\sim 6^\circ$. For the C_{60} molecule, the bending angle is 42° , which is much larger than for these tubes. The C–C bond energy calculated for the C_{60} molecule is smaller than that of graphite¹¹, suggesting that bending the hexagons in C_{60} lowers the bond energy. A similar effect of the bending on bonding energies might apply here. One of the key questions about the tubular structure is how the ABAB hexagonal stacking sequence found in graphite is relaxed, as it is impossible to retain this ideal graphite structure for coaxial tubes. There should be a shortage of 8–9 hexagons in going from one circumference of a tube to that inside it. Disordered graphitic stacking is known as turbostratic stacking, but no detailed accounts of stacking patterns in such structures have been reported. The argument here is also applicable to the spherical graphitic particles mentioned earlier⁶.

All the electron diffraction patterns (Fig. 3) taken from individual carbon needles are indexed by the $\{h0l\}$ and $\{hk0\}$ spots for hexagonal symmetry. The patterns always show strong (00 l) spots when the needle axes are perpendicular to the [001] axis, supporting the idea of a coaxial arrangement of graphitic tubes. As shown in Fig. 2, two side portions of each tube (indicated by shading and labelled 'V') will be oriented so that the

(002) planes satisfy the Bragg diffraction condition for the incident electron beam, and thus give (001)-type spots. Individual (002) planes in these portions are directly imaged in Fig. 1.

The $\{hk0\}$ patterns as a whole show $mm2$ mirror symmetry about the needle axis, and consist of multiple sets of $\{hk0\}$ spots. For example, three sets of $\{hk0\}$ spots seem to form ring patterns, only (100)- and (220)-type spots being seen (Fig. 3a). The diffraction pattern was obtained from a single tube consisting of seven sheets, confirmed by examining the electron microscope image. Referring again to Fig. 2, the top portion of the outermost tube, labelled 'H', and its counterpart on the bottom of the tube, give independently one set each of $\{hk0\}$ patterns. If these two portions of the cylinder have the same orientation, they produce an identical $\{hk0\}$ pattern. If three hexagon sheets on the tubes 1, 2 and 3 were oriented differently, they would give six different $\{hk0\}$ spot patterns. Such a top-bottom effect is one of the requirements for the mirror symmetry. Another requirement is a helical arrangement of carbon hexagons on individual tubes, described further below.

Consider rotation of individual graphite sheets with respect to the needle axes. To explain the graphitic tube structure, the tube is cut at one side along the needle axis and unrolled. This is illustrated schematically in Fig. 4a. Fewer hexagons are drawn, than would form a real tube, but the essential needle geometry correctly represents one of our experimental diffraction patterns. A cylindrical tube can be formed by rolling up the hexagonal sheet about the filament axis (drawn by the heavy line) so as to superimpose hexagons labelled A and B at the top on A' and

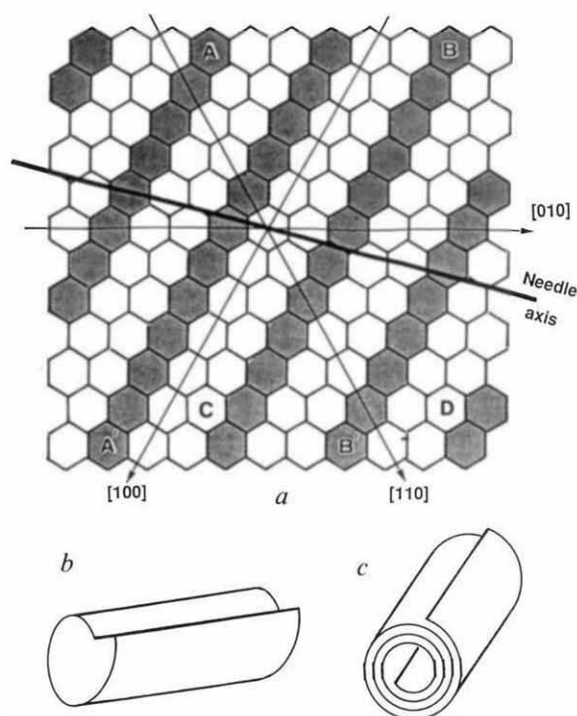


FIG. 4 a, Schematic diagram showing a helical arrangement of a graphitic carbon tubule, which is unrolled for the purposes of explanation. The tube axis is indicated by the heavy line and the hexagons labelled A and B, and A' and B', are superimposed to form the tube (for the significance of C and D, see text). b, The row of hatched hexagons forms a helix on the tube. The number of hexagons does not represent a real tube size, but the orientation is correct. c, A model of a scroll-type filament.

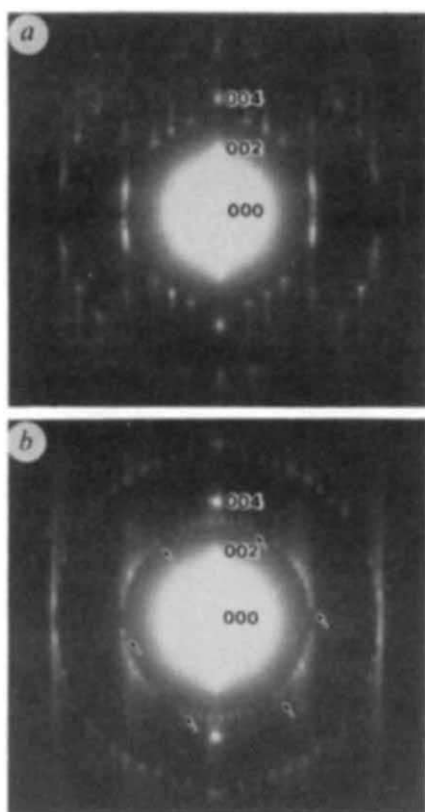


FIG. 3 Electron diffraction patterns from individual microtubules of graphitic carbon. The patterns show $mm2$ symmetry, and are indexed by multiple superpositions of $\{h0l\}$ -type reflections and $\{hk0\}$ reflections of graphite crystal. The needle axes are horizontal. a, Superposition of three sets of $\{hk0\}$ spots taken from a seven-sheet tubule. b, Superposition of four sets of $\{hk0\}$ spots from a nine-sheet tubule.

B' at the bottom. It will be found that the hatched hexagons are aligned perfectly to make a helix around the needle axis (see Fig. 4b). The helical arrangement of the hexagons is responsible for the mirror symmetry as observed in the experimental diffraction patterns. One complete spiral rotation leaves a pitch three hexagons in height at the tip of the needle. There are many possible pitches in the helix, depending on the orientation of the sheet with respect to the needle axis. In other words, the orientation of a hexagon sheet can be determined uniquely by referring to the needle axis. If hexagons A and B coincide with those labelled C and D, the needle axis will be along [010] and thus there will be no spiral rows of hexagons. This has, however, rarely been observed.

Because of the helical structure, the $\{hk0\}$ spot patterns should always contain sets of even numbers of spots. Occasionally, sets of odd numbers of spots occur, such as the groups of three shown in Fig. 3a. These can be explained by frequent coincidence of the top and bottom sheet orientations. Such an accidental coincidence will be increased for hexagonal symmetry of individual hexagon sheet when it is rolled. Consider the sets of four of $\{hk0\}$ in Fig. 3b. Each of the sets of spots, which have equal intensity distributions, is rotated by $\sim 6^\circ$ about the (000) origin, or the needle axis. We take one set of $\{hk0\}$ spots (indicated by arrows), corresponding to one of the hexagon sheets, as a reference whose [100] axis is rotated by 3° about the needle axis. The other three sets of spots are then rotated by 6° , 12° and 24° from the reference sheet. Considering the fact that the needle consists of only nine sheets and each set of $\{hk0\}$ spots is generated from only three or four sheets, it is reasonable that every three or four sheets are rotated stepwise by 6° about the c axis. Any translational shift of the sheets cannot be detected in the electron diffraction patterns. A systematic change in sheet orientations was confirmed by $\{h0l\}$ -type lattice image observations. The question of whether the systematic variation in the

rotation angles in successive tubules acts to stabilize the structure remains to be answered.

The tips of the needles are usually closed by caps that are curved, polygonal or cone-shaped. The last of these have specific opening angles of about 19° or 40°, which can be rationalized in terms of the way that a perfect, continuous hexagonal network can close on itself.

According to Bacon's scroll model for tubular needle growth, needles could be formed by rolling up single carbon-hexagon sheets to form tubular filaments as illustrated in Fig. 4c. Such filaments should have edge overlaps on their surfaces. But I have observed no overlapping edges for the needles described here. Instead, I have observed concentric atomic steps around the needles (I have not confirmed that these are helical). On the basis of these new experimental findings on needle morphologies, I propose a new growth model for the tubular needles. That is, individual tubes themselves can have spiral growth steps at the tube ends (Fig. 4b). It is worth mentioning that the spiral growth steps, which are determined by individual hexagon sheets, will have a handedness. The growth mechanism seems to follow a screw dislocation model analogous to that developed for conventional crystals, but the helical structure is entirely different from the screw dislocation in the sense that the present crystals have a cylindrical lattice. □

Received 27 August; accepted 21 October 1991.

1. Kroto, H. W., Heath, J. R., O'Brien, S. C., Curl, R. F. & Smalley, R. E. *Nature* **318**, 162–163 (1985).
2. Oberlin, A. & M. Endo *J. Cryst. Growth* **32**, 335–349 (1976).
3. Speck, J. S., Endo, M. & Dresselhaus, M. S. *J. Cryst. Growth* **94**, 834–848 (1989).
4. Tibbetts, G. G. *J. Cryst. Growth* **66**, 632–638 (1984).
5. Bacon, R. *J. appl. Phys.* **31**, 283–290 (1960).
6. Iijima, S. *J. Cryst. Growth* **50**, 675–683 (1980).
7. Iijima, S. *J. phys. Chem.* **91**, 3466–3467 (1987).
8. Kroto, H. W. *Science* **242**, 1139–1145 (1988).
9. Krättschmer, W., Lamb, L. D., Fostiropoulos, K. & Huffman, D. R. *Nature* **347**, 354–358 (1990).
10. Whittaker, E. J. W. *Acta Cryst.* **21**, 461–466 (1966).
11. Saito, S. & Oshiyama, A. *Phys. Rev. Lett.* **66**, 2637–2640 (1991).

ACKNOWLEDGEMENTS. I thank Y. Ando for the carbon specimens.

Evidence from Antarctic ice cores for recent increases in snow accumulation

V. I. Morgan*, I. D. Goodwin*, D. M. Etheridge† & C. W. Wooley*

* Australian Antarctic Division, Channel Highway, Kingston 7050, Australia

† CSIRO, Division of Atmospheric Research, Mordialloc, Victoria 3195, Australia

LARGE uncertainties exist in the present knowledge of the mass budget of the Antarctic ice sheet, because of a lack of data on the rates of both ice outflow and snow accumulation¹. Present estimates indicate that both the outflow and the net accumulation are approximately equal to 2,000 km³ of ice per year (equivalent to about 6 mm of sea level)². The temporal variation of accumulation rate is central to determinations of the mass budget, because accumulation can change rapidly in response to short-term climate variations, whereas ice flow varies only on longer timescales. Here we present time series showing changes in the net rate of snow accumulation since 1806 along a 700-km segment of East Antarctica. The accumulation record was derived from the thicknesses of annual layers in ice cores, deduced from seasonal variations in oxygen isotope ratio and in ice-crust stratigraphy. We find a significant increase in the accumulation rate following a minimum around 1960, leading to recent rates that are about 20% above the long-term mean. If this recent increase is widespread, as suggested by shorter-term accumulation data from across a large

TABLE 1 Accumulation data for Wilkes Land sites

Site	Latitude	Longitude	Elevation (m)	Accumulation rate (kg m ⁻² yr ⁻¹)	Accumulation change (%)	
				1975–85	1975–85 to 1955–65	1975–85 to 1930–85
DSS	66° 46' S	112° 59' E	1,370	570	37	19
DE08	66° 43' S	113° 12' E	1,200	1,160	43	28
GD03	69° 00' S	115° 30' E	1,828	355	23	15
GD15	69° 00' S	130° 48' E	2,155	356	26	21

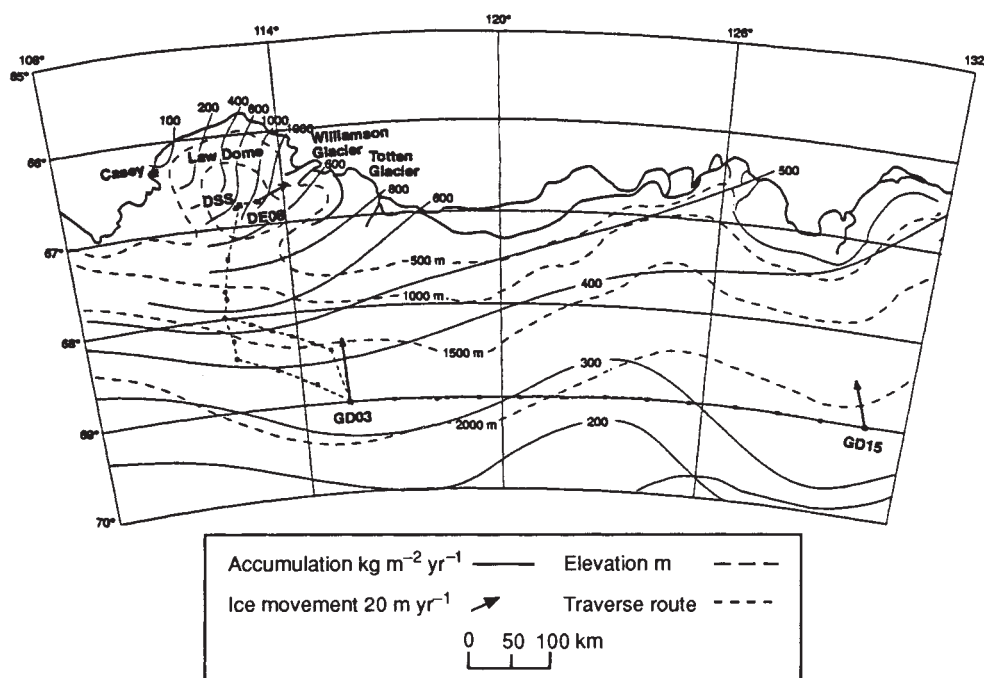
part of Antarctica, the positive imbalance (5–25% of the mass input) shown in recent studies of the ice sheet's mass budget¹ may have existed only since the late 1960s. We estimate that this increase in accumulation rate should contribute to a lowering of sea level of 1.0–1.2 mm per year.

The records come from four ice cores drilled 100–300 km from the coast in Wilkes Land (Fig. 1). Sites DSS and DE08 are located close together (separation 18 km) near the summit of Law Dome in a high accumulation zone. The recent (1978–1987) mean accumulations for the DSS and DE08 sites are 570 and 1,160 kg m⁻² yr⁻¹, respectively. The other two sites, GD03 and GD15, are located on the East Antarctic inland ice sheet in a moderate accumulation zone with recent (1978–1987) mean accumulations of 355 and 356 kg m⁻² yr⁻¹, respectively³.

The thicknesses of the annual accumulation layers were determined by oxygen-isotope ($\delta^{18}\text{O}$) and ice-crust stratigraphy on the cores. The coastal DSS and DE08 cores show exceptionally well preserved annual cycles of $\delta^{18}\text{O}$ because the annual accumulation rates are high and surface melting rarely occurs. In addition, the mixing of surface snow is minimal as both sites are located near the summit of Law Dome where strong surface winds are infrequent (automatic weather station data; I. Allison, Australian Antarctic Division, personal communication). In parts of the cores where the $\delta^{18}\text{O}$ data are ambiguous, electrical conductivity⁴ and hydrogen peroxide⁵ measurements, which also show annual cycles, were used for clarification. The inland GD03 and GD15 cores are located in a different snow accumulation regime where strong katabatic winds redistribute the precipitation⁶. Despite this mixing, there is an annual cycle in $\delta^{18}\text{O}$, although it is not as clear as in the Law Dome cores. The GD03 and GD15 $\delta^{18}\text{O}$ data were supplemented by visible stratigraphy techniques based on identification of an ice crust and wind slab layer which forms annually in autumn and overlies a summer depth hoar layer. For GD03 and GD15 (but not for DSS and DE08) this stratigraphic interpretation was confirmed by stake height measurements of recent accumulation and a reference horizon determined from gross β -radioactivity profiling.

We converted the thickness data from the cores into annual accumulation rate records by allowing for the densification of snow at depth, the upstream variation in snow accumulation (because the ice at depth originated upstream of the borehole) and the vertical strain due to ice flow⁷. Accurate corrections for the densification of firn in the upper layers were made from detailed density measurements on the cores. The correction for the upstream accumulation gradient is negligible for DSS, GD03 and GD15. For DE08 we used the present-day accumulation gradient to calculate a correction of +5% for the deepest ice, reducing linearly to zero at the top of the core. Thinning due to vertical strain in the ice is significant only in the longer DE08 and DSS records. Corrections were made by using measurements of the horizontal strain rate (either direct measurements of strain or calculated from surface velocities) around the boreholes. The calculated strain rates are close to those required to give ice sheet balance with the long-term mean accumulation rate. The vertical strain corrections for the bottom layers in the cores are +8.5% for DSS and +12% for DE08 and are reduced linearly to zero at the surface.

FIG. 1 Location map of Wilkes Land showing the four ice-core sites. Super-imposed on the map are accumulation contours modified from Giovinetto and Bull²¹ using stake and ice-core data from Australian National Antarctic Research Expeditions fieldwork.



The accumulation rate records are shown in Fig. 2. The good correlation between the four records shows that their variability is not due to local effects at the sites but rather reflects temporal variations in accumulation for the Wilkes Land region. In addition, because the dating of the layers in each core is independent, the correlation confirms that there are no systematic errors in the layer identification and hence in the layer thicknesses. To allow comparison between the sites, the records have been normalized against the mean for the 1930–1985 epoch, which is common to all four cores.

The longest records, from Law Dome, indicate an overall slight increase in accumulation from 1806 to the present, but there is large variability on timescales of up to 100 years. Significant changes in accumulation rate have occurred since 1955. Accumulation decreased suddenly around 1955 to reach a minimum around 1960. Since then, accumulation has steadily increased to the present values, which are the highest in the record. Verification of this trend across Antarctica is difficult as few comparable records exist. Pourchet *et al.*⁸ used measurements of total β -radioactivity to identify the 1955 and 1965 layers for 14 sites in East Antarctica. This allowed average accumulation to be compared between the consecutive epochs 1955–1965 and 1965–77. They found an average increase of 30% from 1955–65 to 1965–77 with increases at two inland sites, South Pole and Dome C, of 10% and 33% respectively. Peel⁹ also reported a large recent accumulation increase of ~20% since 1950 in data from oxygen isotope stratigraphy of Antarctic Peninsula ice cores. To investigate the change in Wilkes Land, we analysed our data in the two common epochs, 1955–65 and 1975–85, for each of the four sites. The results are listed in Table 1 and show that the magnitude of the accumulation increase varies with accumulation rate, with the largest percentage increase occurring within the high-accumulation coastal escarpment region.

Accumulation in Antarctica is controlled principally by air temperature (moisture transport is limited by the saturation vapour pressure of the cold air) and, particularly in the coastal zone, by the frequency of 'precipitation' events, which depends on the intensity of cyclonic activity¹⁰. Empirical studies^{11–13} do show correlation between temperature and accumulation (albeit for geographical rather than temporal variation), but the high accumulation in winter compared with that in summer arises from the increased intensity of cyclonic activity in winter, which overrides the accumulation signal produced by the effect of temperature and moisture transport. Our time-varying accumulation record also shows correlation with mean Antarctic temperatures (data from Jacka¹⁴ in Fig. 2), but this will not be a direct cause and effect because the change in accumulation is much too large (~50% per °C) to be due to temperature. Estimates of the temperature effect range from 3–4% per °C for global average general circulation models¹⁵ up to 10% per °C for polar models⁷. Estimates from the variation in the saturation vapour pressure of water over ice¹⁶ also give ~10% per °C. The

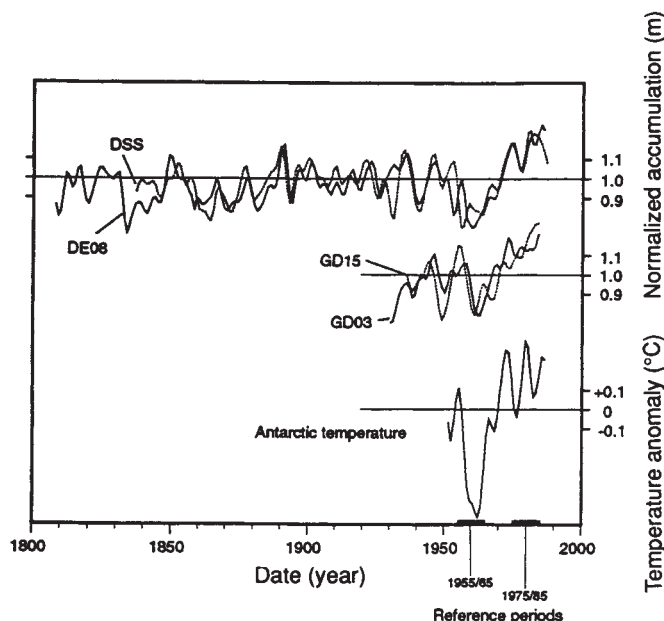


FIG. 2 Accumulation data for the four cores compared with Antarctic temperature records. The corrected accumulation data are normalized by dividing by the mean value from 1930 to 1985 to allow comparison between cores. The data is smoothed by a gaussian filter (bandwidth 3 years) to reduce the large fluctuations in measured accumulation principally caused by surface roughness of the ice cap. The Antarctic temperature results are a compilation of mean annual temperatures from the Antarctic Stations with long, continuous records and Macquarie Island. These data are also smoothed with the same filter.

dominant influence on increased accumulation in Wilkes Land (since 1958, when records began) seems to be the observed intensification of cyclonic activity within the quasi-stationary cyclone situated near Casey Station at 100°–110° E in the circum-polar low-pressure trough (R. Allan, CSIRO-DAR, personal communication). We suggest that a general intensification of cyclonic activity around Antarctica is the main cause of the observed increase in accumulation. The effect is most pronounced near the coast, but even inland precipitation (of ice crystals from clear skies) is influenced by the advection of moist air over the continent by cyclonic activity¹⁰. Because precipitation events are accompanied by higher temperatures¹⁷, changes in cyclonic activity leading to more frequent or longer-lasting events will lead to higher average temperatures and hence the observed correlation between accumulation and temperature.

A recent compilation of mass-balance data for Antarctica¹ concludes that average accumulation is in excess of outflow by 5–25% (or 17–24% for the total grounded ice area minus the Antarctic Peninsula). A large part of the accumulation data, especially for the high-accumulation coastal areas, has been collected or updated since 1970. If these values are up to 20% above the mean for the last 200 years, as our data suggest, this implies that before the recent increase, the ice sheet was close to balance. An alternative method of estimating the state of balance of the ice sheet is by direct topographic measurement. So far, surface surveys using satellite geodetic positioning techniques at Law Dome summit indicate no large elevation change; the total measurement error, however, is ~2 m, which is comparable with the expected height change over the 10-year measurement period¹⁸.

Independent of possible causes, a direct and instantaneous result of accumulation increase is a lowering of sea level. If we assume that the ice sheet was in balance with the mean accumulation over the past 170 years and that the accumulation variations seen in Wilkes Land are widespread, we can estimate an equivalent sea-level lowering due to the recent mass-balance change. We assumed that the Wilkes Land accumulation increase (23% coastal and 18% interior zones) was directly proportional to that over the total Antarctic grounded ice area. The corresponding total areas for the coastal escarpment and interior were calculated¹¹ from SPRI maps. We based our calculations on the currently accepted range¹⁹ for surface mass input of $1,817$ to $2,168 \times 10^{12} \text{ kg yr}^{-1}$. Accordingly, we calculated a negative contribution to sea level of 1.0 – 1.2 mm yr^{-1} . This is slightly greater than the $0.75 \pm 0.25 \text{ mm yr}^{-1}$ reported by Meier²⁰ from mass budget calculations that include large uncertainties in ice flow. □

Isotopic evidence for involvement of CO₂-bearing magmas in granulite formation

James Farquhar & Thomas Chacko

Department of Geology, University of Alberta, Edmonton, Alberta T6G 2E3, Canada

It has been suggested^{1,2} that CO₂-bearing magmas play an important part in the formation of granulites. These magmas crystallize directly into granulites, and also expel fluids which promote the development of granulite-facies mineral assemblages in adjacent country rocks. Direct evidence for carbonic fluids remains elusive, however, as granulite-facies aureoles adjacent to charnockitic intrusives¹ may result either from the influx of carbonic fluids derived from the intrusives or simply from extraction of water into a vapour-absent melt. Here we report results of a detailed study of carbon isotope compositions of graphite associated with a charnockite dyke from the Ponmudi quarry of south India. Higher graphite abundances and an anomalously heavy carbon isotopic composition ($\delta^{13}\text{C}_{\text{gr}} = -8.1\%$) at the dyke margins indicate that externally derived CO₂-rich fluids were transported into the Ponmudi paragneisses by the intrusive, and precipitated nearly quantitatively as graphite on encountering reducing country rocks. The occurrence of large orthopyroxene crystals along the dyke margins suggests that these fluids also served to dehydrate the country rocks. This dyke may thus be representative of the widespread transport of carbonic fluids in this terrain by felsic magmas.

The Ponmudi quarry is located near the southern tip of India within the Kerala Khondalite belt, a region of high-grade meta-sedimentary rocks consisting primarily of pelitic and semipelitic paragneisses along with subordinate mafic granulites, calc-silicate rocks and quartzites^{4–6}. The quarry consists of well-banded garnet-biotite paragneiss which is partially overprinted by patches of coarse-grained charnockite⁷. The charnockitic alteration is relatively diffuse and nearly isochemical, features which have led several workers to conclude that incipient charnockite formation resulted from the passage of carbonic fluids^{4–9}. Jackson *et al.*¹⁰ provide compelling evidence for externally derived CO₂ in the quarry by documenting positive shifts in both $\delta^{13}\text{C}_{\text{CO}_2}$ and $\delta^{13}\text{C}_{\text{gr}}$ (defined in Fig. 1) for charnockite patches relative to the paragneiss. The transporting agent that brought this CO₂ into the quarry is unknown.

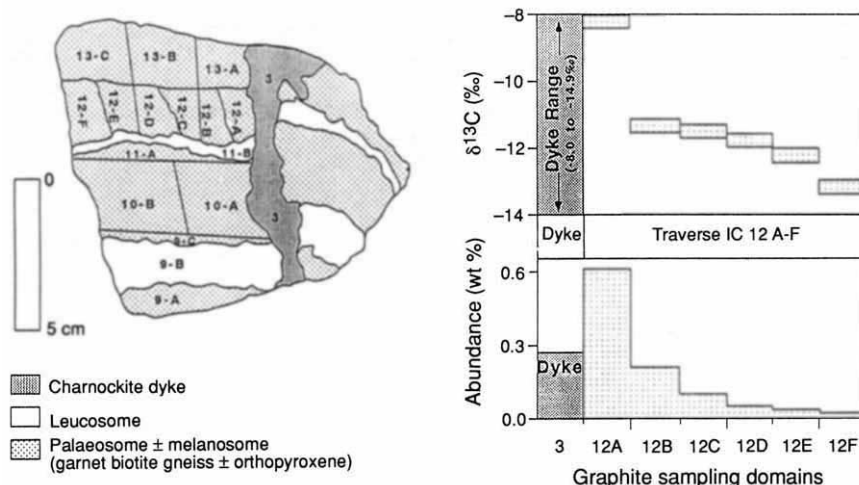
At the eastern end of the Ponmudi quarry, a narrow charnockite dyke (2–7 cm wide) cuts across and locally displaces gneissic banding. The dyke, which is more felsic than the country rocks, consists dominantly of K-feldspar (~60%) and quartz (~25%) along with small amounts of high-titanium biotite (~5%), plagioclase (~5%), and minor (<5%) garnet, chloritized orthopyroxene and calcite. The most intense formation of charnockite and the highest concentrations of graphite in the quarry occur along the dyke margins, making it a natural focus for examining possible CO₂ sources accompanying orthopyroxene formation. Throughout the quarry, graphite occurs as finely disseminated flakes (1 mm) in both the paragneiss and the charnockitic patches. Within and along the dyke margins, graphite occurs as coarser-grained (2–4 mm) hexagonal flakes, which are commonly in contact with biotite. X-ray diffraction patterns indicate a fully ordered graphite ($d_{002} = 3.357$, full width at half maximum $0.18^\circ(2\theta)$) characteristic of those found in high-temperature environments¹¹.

Graphite abundances remain constant throughout the paragneiss and charnockite at $\sim 0.02 \pm 0.01 \text{ wt\%}$ but increase sharply to $0.61 \text{ wt\%}$ near the dyke. This increase is accompanied by a large shift in $\delta^{13}\text{C}_{\text{gr}}$ from -19.3% in the paragneiss to -8.1% immediately adjacent to the dyke (Fig. 1), strongly suggesting that the dyke introduced externally derived carbonic fluids into

Received 3 June; accepted 12 September 1991.

- Bentley, C. R. & Giovinetto, M. B. in *Proc. Int. Conf. on the Role of Polar Regions in Global Change, Fairbanks, Alaska, June 1990* (Arctic Research Consortium of the United States, in the press).
- Budd, W. F. & Summonds, I. in *Proc. Int. Conf. on the Role of Polar Regions in Global Change, Fairbanks, Alaska, June 1990* (Arctic Research Consortium of the United States, in the press).
- Goodwin, I. D. *Antarct. Sci.* **2**, 235–242 (1990).
- Hammer, C. U. *J. Glaciol.* **25**, 359–372 (1980).
- Sigg, A. & Nettel, A. *Ann. Glaciol.* **10**, 157–162 (1988).
- Goodwin, I. D. *J. Glaciol.* (in the press).
- Patterson, W. S. B. & Waddington, E. D. *Rev. Geophys. Space Phys.* **22**, 123–130 (1984).
- Pourchet, M., Pinglot, F. & Lorus, C. *J. geophys. Res.* **88**, 6013–6020 (1983).
- Peel, D. A. in *Climate since 1500 AD* (eds Bradley, R. S. & Jones, P. D.) (Harper Collins Academic, London, in the press).
- Bromwich, D. H. *Rev. Geophys.* **26**, 149–168 (1988).
- Fortuin, J. P. F. & Oerlemans, J. *Ann. Glaciol.* **14**, 78–84 (1990).
- Loewe, F. *J. geophys. Res.* **67**, 5171–5177 (1962).
- Mellor, M. *Polarforschung* **5**, 179–180 (1963).
- Jacka, T. H. *Ann. Glaciol.* **14**, 127–130 (1990).
- Mitchell, J. F. B., Manabe, V., Meleshko, T. & Tokioka, T. in *Climate Change. The IPCC Scientific Assessment* (eds Houghton, J. T. *et al.*) (Cambridge University Press, 1990).
- Robin, G. de Q. *Phil. Trans. R. Soc. B280*, 143–168 (1977).
- Loewe, F. *Z. Gletscherkunde Glazialgeol.* **10**, 189–197 (1974).
- Young, N., Malcolm, P., Mantell, P. & McGibbon, E. *Ann. Glaciol.* **12**, 220 (1989).
- Warrick, R. & Oerlemans, J. in *Climate Change. The IPCC Scientific Assessment* (eds Houghton, J. T. *et al.*) (Cambridge University Press, 1990).
- Meier, M. F. *Nature* **343**, 115–116 (1990).
- Giovinetto, M. B. & Bull, C. *Byrd Polar Research Center Rep. No. 1*, 49 (1987).

FIG. 1 Illustration of sample IC. The sample is divided into three rock types, the charnockite dyke, palaeosome and leucosome. The melanosome generally occupies the 0.5 cm between the leucosome and the palaeosome. Inset: Plots of abundance (wt%) and isotopic composition (‰) against position for sampling traverse IC 12 A–F. The dyke sample IC 3 is represented by a darker stipple and the range of $\delta^{13}\text{C}_{\text{gr}}$ values. Uncertainties for isotopic determinations are represented by the thickness of the boxes. $\delta^{13}\text{C} = [({}^{13}\text{C}/{}^{12}\text{C})_{\text{sample}}/({}^{13}\text{C}/{}^{12}\text{C})_{\text{standard}} - 1]$, multiplied by 1,000 to give values in ‰. Estimates of minimum graphite abundance were made by comparing masses for graphite separates with whole rock sample masses. Reproduction of individual measurements yielded uncertainties better than $\pm 0.01\%$.



the host rocks. Moreover, the occurrence of large orthopyroxene grains along the dyke margins indicates that intrusion occurred under granulite-facies conditions. We propose, therefore, that the dyke represents a CO_2 -bearing magma which expelled fluids into reduced country rocks during crystallization, resulting in both graphite precipitation and charnockite formation along the dyke margins.

The variations in graphite abundances and isotopic composition along the dyke margins result primarily from mixing of isotopically light graphite, which was originally in the paragneiss, with a second population of heavy graphite precipitated from incoming CO_2 fluid. Best fits to the data are achieved if the precipitated graphite has $\delta^{13}\text{C} = -9.0$ to -6.5% (Fig. 2). Deviations from simple endmember mixing curves result from variable initial graphite abundances and possibly also partial re-equilibration of graphite with incoming CO_2 . We note, however, that slow rates for carbon diffusion in graphite^{12–14} combined with slow rates for graphite recrystallization^{13,15} inhibit extensive re-equilibration of isotopes.

Even without complete re-equilibration, an estimate of $\delta^{13}\text{C}_{\text{CO}_2}$ of the fluid from which graphite precipitated can be made once the fraction of fluid consumed by precipitation is known. Lamb and Valley¹⁶ demonstrated that carbonic fluids entering reduced rocks must precipitate graphite on a quantitative basis until the oxygen fugacity of the rock is high enough to be in equilibrium with a CO_2 -rich fluid. For quantitative or near-quantitative precipitation in a closed system (70–100% of original fluid precipitated), our data from the dyke margin require precipitation from CO_2 fluids with initial $\delta^{13}\text{C}_{\text{CO}_2}$ between -8 and -4% (Fig. 3). Precipitation of graphite in the less-reducing environment of the CO_2 -bearing dyke bridges a continuum between open and closed systems. This is reflected both by a distinctly lower abundance of graphite within the dyke (0.27 wt%) relative to immediately adjacent host rocks (0.61 wt%) and by variable $\delta^{13}\text{C}_{\text{gr}}$ values within the dyke (-8.0 to -14.9% ; mean of nine samples -11.1%). It is the interplay between the relative rates of fluid introduction and graphite precipitation that results in this continuum. We suggest, therefore, that small-scale variation of $\delta^{13}\text{C}_{\text{gr}}$ from the dyke stems from a local control of these rates. We do not favour simple endmember mixing models for the dyke, as they do not account for the relatively high graphite abundances and the large range of $\delta^{13}\text{C}_{\text{gr}}$ values.

The range of $\delta^{13}\text{C}_{\text{CO}_2}$ (-8 to -4%) corresponds closely to the range for mantle-derived CO_2 fluids^{17,18} but cannot unambiguously be distinguished from CO_2 derived from extensively decarbonated sedimentary carbonates or appropriate mixtures of carbonate and biogenic carbon¹⁹. The paucity of carbonate lithologies in the Khondalite belt ($<1\%$) makes the latter sources

less likely. Carbon dioxide of this composition may also be generated internally by repeated mobilization and redeposition of organic carbon (incremental batch oxidation and quantitative reduction)²⁰. Mass balance dictates, however, that only small quantities of ^{13}C -enriched graphite are produced from a given reservoir of organic material; this contrasts with the observations at Ponnudi, where the highest concentrations of graphite are also the most enriched in ^{13}C . The range of $\delta^{13}\text{C}_{\text{CO}_2}$ values calculated above is similar to those measured in fluid inclusions from charnockite patches at Ponnudi and other localities in the Khondalite Belt^{10,21}, suggesting a connection between the fluids associated with the charnockite dyke and those responsible for the development of incipient charnockites in the terrain. Thus,

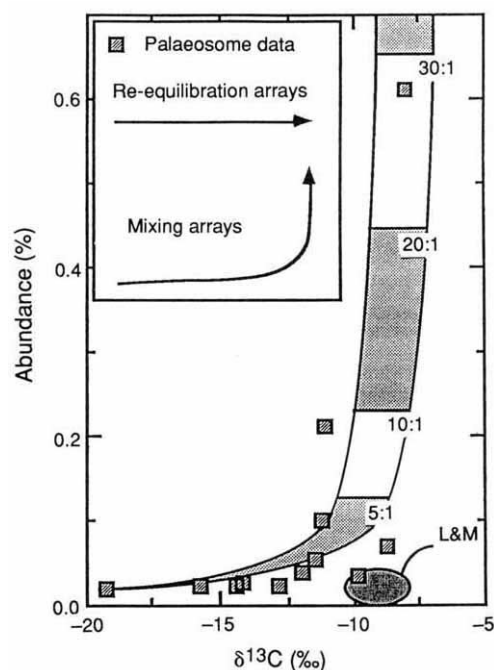


FIG. 2 Data for the palaeosome plotted as abundance against $\delta^{13}\text{C}$. Mixing curves are constructed for endmembers with initial graphite of 0.02 wt%, $\delta^{13}\text{C} = -19.3\%$ and added graphite of $\delta^{13}\text{C} = -9.0\%$ and -6.5% . On this diagram, re-equilibration occurs along horizontal lines (constant abundance), and mixing occurs along curved lines (isotopic composition function of abundance). Uncertainties are represented by the size of data boxes. Data for the leucosome and melanosome plot in a separate field (L&M) because of their different initial abundances of graphite.

on a small scale, the isotope data provide a link between deep-seated carbonic fluids, felsic magmas and the formation of granulites.

The volume of rock found at present within the dyke at Ponmudi is minor, and therefore the potential for this small volume of magma to transform host rocks to charnockite is correspondingly small ($\sim 20\%$ of the volume of the dyke)¹. If, however, the dyke is considered as a conduit for magma transport rather than as a static magma, then a much larger volume of magma may have passed through the quarry by way of this relatively narrow dyke. For example, given a magma flow velocity in the dyke of $\sim 200 \text{ m yr}^{-1}$ (equation 5 in ref. 22), all the charnockitization in the quarry can be accounted for by a few hundred years of magma flow. This hypothesis is in keeping with a recent proposal²³ that transport of large granitic batholiths from lower crustal source regions to upper crustal zones of emplacement occurs primarily by upward passage of melts through a network of narrow dykes. If these dykes are also fluid-bearing, as at Ponmudi, then they must be considered as a potentially important transporting agent for fluids.

The Khondalite belt was widely intruded by felsic pegmatite dykes at 500–600 Myr, roughly contemporaneous with granulite facies metamorphism^{5,24,25}. Many of these dykes are graphite-bearing with reported $\delta^{13}\text{C}_{\text{gr}}$ values from -10 to -13% (refs 26, 27). These values are similar to those found in the dyke at

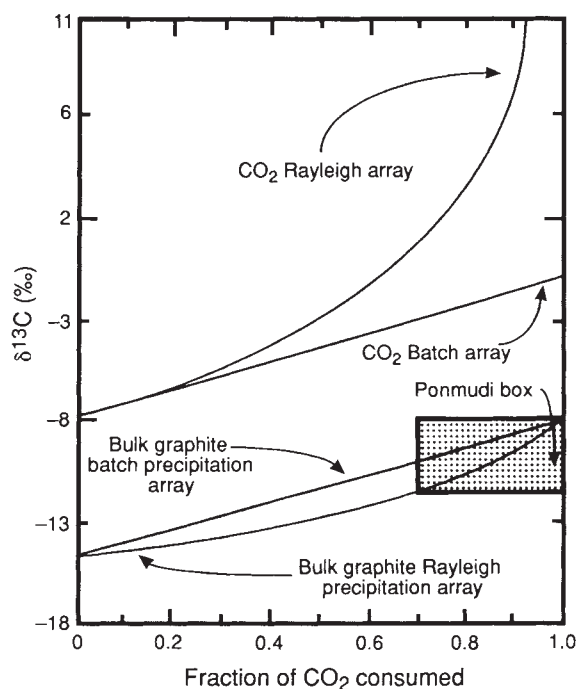


FIG. 3 Plot illustrating the effect of Rayleigh and batch precipitation of graphite from CO_2 -fluids with initial $\delta^{13}\text{C} = -8.0\%$ at 750°C ($\Delta^{13}\text{C}_{\text{CO}_2\text{-graphite}} = 6.8\%$)¹⁶. The upper two lines represent compositions traced out by CO_2 -fluids undergoing closed-system batch and Rayleigh precipitation of graphite. The lower two arrays represent isotopic variation for graphite resulting from closed-system batch and Rayleigh precipitation. The Ponmudi box represents the range of expected carbon isotopic compositions for precipitation of 70–100% of initial CO_2 to form graphite. The necessary condition for batch precipitation is that the rates of graphite precipitation exceed the rates of introduction for CO_2 -fluids. Rayleigh precipitation requires rates for carbon isotopic equilibration that are slower than rates for graphite precipitation (that is, in this environment, combination of slow exchange rates^{12–14} and high precipitation rates³⁵ can produce isotopically equilibrated rims on graphite crystals which inhibit further exchange between the interiors of these crystals and isotopically changing carbonic fluids.) Consequently, graphite precipitated by batch precipitation is expected to show homogenous $\delta^{13}\text{C}$, whereas graphite precipitated by Rayleigh precipitation is expected to show $\delta^{13}\text{C}$ zonation (^{13}C increases from core to rim).

Ponmudi but distinctly heavier than disseminated graphite from typical paragneisses. We propose that the dykes represent the vestiges of an extensive conduit network which allowed relatively large amounts of magma and deep-seated carbonic fluids to pass through the terrain. Fluids expelled from these magmas infiltrated the immediately adjacent wall rocks and were also strongly channelled along zones of weakness over a much broader area, as indicated by ^{13}C enrichments of 2% along several leucosome–melanosome boundaries, both adjacent to and at some distance from the dyke at Ponmudi.

The effect of carbonic fluid influx on country rocks is determined by the temperature of infiltration. In some cases, such as at Ponmudi, magma and fluids were introduced near peak metamorphic conditions and orthopyroxene formation resulted⁶. In many other cases in the Khondalite belt, interaction of magma and fluids with host rocks persisted to lower temperature conditions and retrogressed previously formed high-grade assemblages⁶. The predominance of retrograde relationships can be attributed to a preservational bias for late-stage features in a dynamic flow system that was active throughout the entire metamorphic event. Thus, preservation of prograde relationships is expected only in cases where the supply of magma to the dyke was interrupted at high-temperature conditions. In some situations, complete evacuation of magma from the fracture may have occurred; subsequent closing and healing of such evacuated fractures would make the evidence for the passage of melts very cryptic²³.

These dykes may be part of a larger system that extends from the mantle to the upper crust. Influx of mantle-derived magmas and carbonic fluids into the southern Indian shield at 750–550 Myr (refs 28–30) promoted melting in the lower crust. These melts moved through mid-crustal levels in an extensive network of dykes and (cryptic) conduits to the upper crust, transporting both heat and fluids to higher levels. Evolution of CO_2 -rich fluids from magmas during ascent and subsequent channelling of these fluids into host rocks resulted in patchy granulite formation. Interaction of fluids with reducing country rocks also caused nearly quantitative precipitation of graphite with mantle-like isotopic compositions. Economic-grade graphite deposits in Kerala and neighbouring Sri Lanka ($\delta^{13}\text{C}_{\text{gr}} = -7.8 \pm 2\%$)³¹ may represent higher-level or later-stage manifestations of the same magma–fluid system^{32,33}. The net volume of material transported by this dyke network greatly exceeds the present volume of the dykes and permitted transport of large quantities of CO_2 through this part of Gondwanaland. Transport of carbonic fluids by similar dyke networks may be important in other terrains where large-scale influx of mantle-derived CO_2 is suspected. Evidence for these fluids should be sought at contacts between dykes and relatively reducing metasedimentary country rocks where graphite precipitation is most likely to occur. □

Received 28 May; accepted 3 September 1991.

1. Frost, B. R. & Frost, C. D. *Nature* **327**, 503–506 (1987).
2. Peterson, J. W. & Newton, R. C. *Nature* **340**, 378–380 (1989).
3. Waters, D. J. *J. metamorphic Geol.* **6**, 387–404 (1988).
4. Hansen, E. C., Janardhan, A. S., Newton, R. C., Prame, W. K. B. N. & Ravindra Kumar, G. R. *Contrib. Miner. Petrol.* **96**, 225–244 (1987).
5. Srikantappa, C., Raith, M. & Spiering, B. *J. geol. Soc. India* **26**, 849–872 (1985).
6. Ravindra Kumar, G. R. & Chacko, T. *J. geol. Soc. India* **28**, 277–288 (1986).
7. Ravindra Kumar, G. R., Srikantappa, C. & Hansen, E. *Nature* **313**, 207–209 (1985).
8. Raith, M., Hoernes, S., Klatt, E. & Stahle, H. J. in *Fluid Movements—Element Transport and the Composition of the Deep Crust* (ed. Bridgewater, D.) 29–38 (Kluwer, Amsterdam, 1989).
9. Harris, N. B. W. & Bickle, M. J. *Earth planet. Sci. Lett.* **93**, 151–156 (1989).
10. Jackson, D. H., Matthey, D. P. & Harris, N. B. W. *Nature* **333**, 167–170 (1988).
11. Grew, E. S. *J. Geol.* **82**, 50–73 (1974).
12. Valley, J. W. & O'Neill, J. R. *Geochim. cosmochim. Acta* **45**, 411–419 (1981).
13. Chacko, T., Mayeda, T. K., Clayton, R. N. & Goldsmith, J. R. *Geochim. cosmochim. Acta* (in the press).
14. Wada, H. *Nature* **331**, 61–63 (1988).
15. Fischbach, D. B. *Chemistry and Physics of Carbon* Vol. 7 (ed. Walker, P. L.) 1–107 (1971).
16. Lamb, W. M. & Valley, J. W. *Nature* **312**, 56–58 (1984).
17. Taylor, B. E. *Miner. Soc. Am. Rev. Min.* **16**, 185–226 (1986).
18. Matthey, D. P. *Terra Cogn.* **7**, 31–37 (1987).
19. Valley, J. W. *Miner. Soc. Am. Rev. Min.* **16**, 445–490 (1986).
20. Hoernes, S. & Raith, M. *Fortschritte Min.* **66**, 65 (1988).
21. Santosh, M., Harris, N. B. W., Jackson, D. H. & Matthey, D. P. *J. Geol.* **98**, 915–926 (1990).

22. Turcotte, D. L. in *Magmatic Processes: Physicochemical Principles* (ed. Mysen, B. O.) *Geochem. Soc. spec. Publ.* **1**, 69–74 (1987).
23. Clemens, J. D. & Mawer, C. K. in *Limpopo Field Workshop* (ed. Barton, J. M. Jr), 32–33 (1990).
24. Buhl, D. thesis, Univ. of Munster (1987).
25. Soman, K. & Nair, N. G. K. *Min. Mag.* **49**, 733–738 (1985).
26. Soman, K., Lobzova, R. V. & Sivasdas, K. M. *Econ. Geol.* **81**, 997–1002 (1986).
27. Klatt, E., Hoernes, S. & Raith, M. *J. geol. Soc. India* **31**, 57–59 (1988).
28. Santosh, M. & Drury, S. A. *J. Geol.* **96**, 616–626 (1988).
29. Deans, T. & Powell, J. L. *Nature* **218**, 750–752 (1968).
30. Hansen, E. C., Hickman, M. H., Grant, N. K. & Newton, R. C. *EOS* **66**, 419–420 (1985).
31. Dobner, A., Graf, W., Hahn-Weinheimer, P. & Hirner, A. *Lithos* **11**, 251–255 (1978).
32. Katz, M. B. *Mineral. Deposita* **22**, 18–25 (1987).
33. Silva, K. K. M. W. *Econ. Geol.* **82**, 1710–1722 (1987).
34. Bottinga, Y. *Geochim. cosmochim. Acta* **3**, 49–64 (1969).
35. Holloway, J. R. & Reese, R. L. *Am. Miner.* **59**, 587–597 (1974).

ACKNOWLEDGEMENTS. We thank R. W. Luth, K. Muehlenbachs, B. E. Nesbitt, J. Steer and the stable isotope group at the University of Alberta for assistance and discussions. We also thank NSERC for financial support and K. Muehlenbachs for use of the stable isotope facilities. Finally, we thank S. Hoernes for comments and clarifications to the manuscript.

Conservation of polymorphic simple sequence loci in cetacean species

Christian Schlötterer*, Bill Amos† & Diethard Tautz*

Institut für Genetik und Mikrobiologie der Universität München, Maria-Wardstrasse 1a, 8000 München 19, Germany

† University of Cambridge, Department of Genetics, Downing Street, Cambridge CB2 3EH, UK

LENGTH polymorphisms within simple-sequence loci occur ubiquitously in non-coding eukaryotic DNA and can be highly informative in the analysis of natural populations^{1–4}. Simple-sequence length polymorphisms (SSLP) in the long-finned pilot whale *Globicephala melas* (Delphinidae) have provided useful information on the mating system as well as on the genetic structure of populations⁵. We have therefore tested whether the polymerase chain reaction primers designed for *Globicephala* could also be used to uncover variability in other whale species. Homologous loci could indeed be amplified from a diverse range of whales, including all toothed (Odontoceti) and baleen whales (Mysticeti) tested. Cloning and sequencing these loci from 11 different species revealed an unusually high conservation of sequences flanking the simple-sequence stretches, averaging 3.2% difference over 35–40 Myr. This represents the lowest divergence rate for neutral nucleotide positions found for any species group so far and raises the possible need for a re-evaluation of the age of the modern whales. On the other hand, the high conservation of non-coding sequences in whales simplifies the application of SSLP DNA fingerprinting in cetacean species, as primers designed for one species will often uncover variability in other species.

We have isolated several simple-sequence loci randomly from the long-finned pilot whale *Globicephala melas* containing either GT/AC or GA/TC dinucleotide repeats. Primers flanking these stretches were synthesized and used to amplify enzymatically the respective loci from several hundred individuals, revealing the expected SSLPs. The frequency distribution of the different alleles varied between independently caught pods, allowing us to determine the genetic differentiation of these pods, and to confirm and extend inferences about the mating system in these whales (ref. 5 and B.A., C.S. and D.T., in preparation).

Primers developed for *Globicephala* were then tested on DNA isolated from representatives of most of the major cetacean radiations (baleen whales, sperm whales, dolphins, porpoises and beaked whales). We found that all loci tested could be amplified in all species tested. Such a level of conservation can be contrasted with the results of a recent study on the conservation of simple-sequence loci between sheep and cow where it

was found that only about 40% of the loci could be amplified in the other species⁶.

To investigate whether equivalent loci in mysticete whales (baleen whales) show the same kind of polymorphisms as they do in the odontocete pilot whale (a toothed whale), four loci were tested on a number of randomly selected, and hence presumably unrelated individuals from two different mysticete whale species, the minke whale (*Balaenoptera acutorostrata*) and the fin whale (*B. physalus*). Figure 1 shows that polymorphism can indeed be detected for all but one locus. To demonstrate true homology, the amplified fragments of these loci were cloned and sequenced from 11 different whale species. In all cases we found the expected simple-sequence stretch. In addition, the flanking sequences showed an unusually high degree of conservation (Fig. 2), the average difference between toothed whales and baleen whales being about 3.2% (Fig. 3).

The average typical divergence rate for neutral nucleotide positions is usually about 0.5% Myr⁻¹ (ref. 7), although the lowest found so far is about 0.2% Myr⁻¹ in the primates⁸. These rates are, however, usually calculated on the basis of third codon positions, which may not be entirely neutral⁹. In contrast, our DNA samples represent randomly cloned pieces of DNA, which are unlikely to be part of coding regions, as the simple sequence stretches and the polymorphism therein would disrupt any open reading frame. On the other hand, the possibility that our loci are derived from potentially conserved regulatory regions cannot be ruled out. However, we consider this unlikely, as all four loci shown here, as well as preliminary data on three further loci (not shown), show a similar divergence pattern. We believe therefore that most of the loci analysed are representative of non-coding neutral DNA (compare Fig. 1 legend).

If one assumes a split between the odontocete and mysticete whales 35–40 Myr ago^{10,11}, one can calculate from our data set a rate of neutral nucleotide substitution of about 0.09% Myr⁻¹, which would represent the lowest found so far for any species

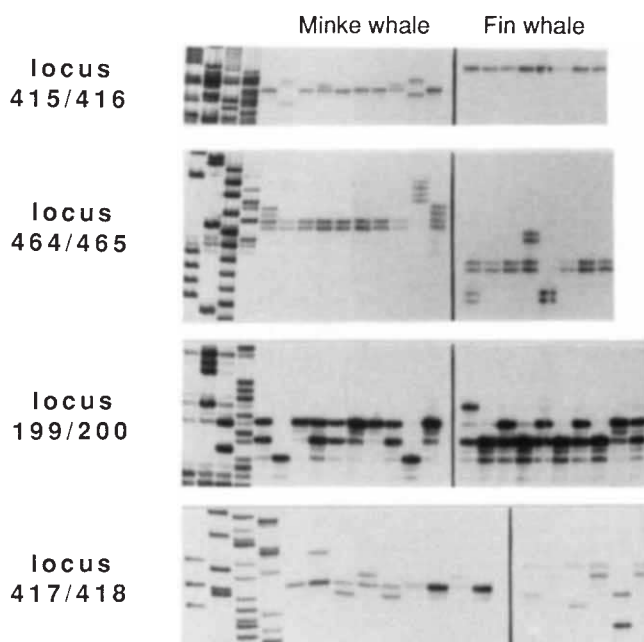


FIG. 1 Test for polymorphism of the four simple-sequence loci in a population sample from minke whales (*B. acutorostrata*) and fin whales (*B. physalus*) according to described methods^{2–4}. Randomly chosen individuals were tested for each locus. The sequencing reaction at the sides serve as size marker. Locus 415/416 in the fin-whale sample shows no polymorphism in the animals we tested. This may be explained by the fact that in this species, there are three single-nucleotide substitutions within the simple-sequence stretch (Fig. 2). It is possible that these substitutions prevent efficient slippage of the remainder of the stretch and therefore reduce the polymorphism of this particular locus.

* Present address: Zoologisches Institut der Universität München, Luisenstrasse 14, 8000 München 2, Germany.

oodon ampullatus), the sei whale (*Balaenoptera borealis*), the fin whale (*Balaenoptera physalus*), the minke whale (*Balaenoptera acutorostrata*) and the humpback whale (*Megaptera novaengliae*). Sequences flanking the simple-sequence stretches were used for a homology search against the nucleic acid data libraries. This search revealed that part of locus 417/418 has some homology with mammalian repetitive L1 elements (reverse complementary strand of the 5'-half with position 5,635–5,700 in the human L1 sequence; Genbank Accession No. M19503). However, this locus behaves as a single-copy mendelian locus in the population analysis, suggesting that it has its 5'-end located in a truncated, non-functional L1 element and its 3'-end in unique flanking DNA.

NATURE • VOL 354 • 7 NOVEMBER 1991

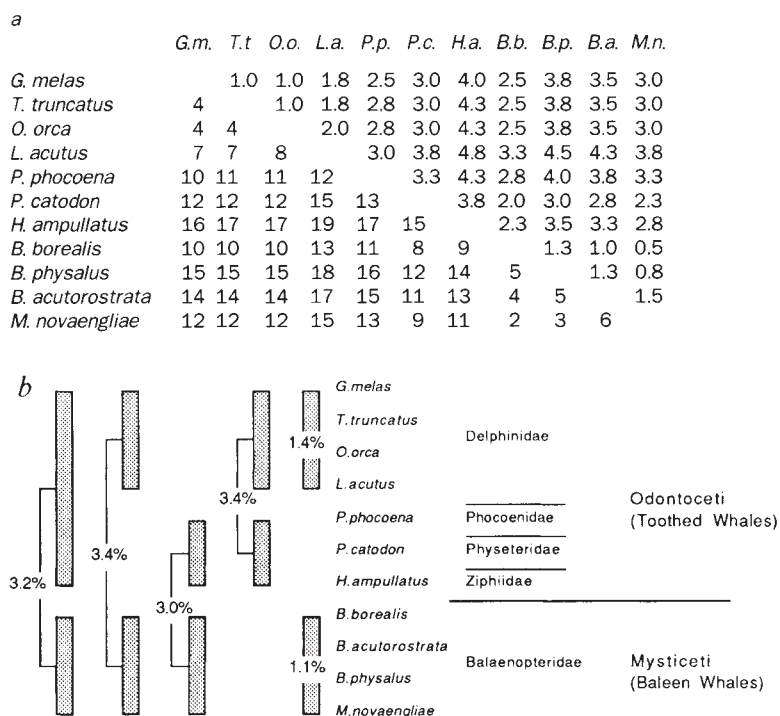


FIG. 3 **a**, Pairwise distance-matrix of the sequences shown in Fig. 2. Lower left, counted differences, upper right: per cent difference values. Insertions/deletions were counted as single events. The simple-sequence stretches (marked with 'x' in Fig. 2) were excluded from this comparison, as differences here are expected to arise by slippage¹⁷ and cannot be compared with point mutations. This leaves a total of 397-nucleotide flanking sequences which were compared between all 11 species. The percentage difference values were calculated without correction for potential back mutations. **b**, Average difference values between the species groups, calculated on the basis of the data in the distance matrix. The major groups show similar difference values (between 3.0 and 3.4%), supporting the inference of an explosive radiation of the modern whales.

group. Interestingly, an apparently lowered mutation rate was also found in a comparison of mitochondrial sequences between minke whales and dolphins¹². Differences in the rate of neutral nucleotide substitutions can be correlated with the generation times of the species groups⁸. Generation times of cetaceans are, however, roughly comparable with those of primates and cannot therefore account for the slower rate. There are two alternative explanations. Either whale DNA evolves slowly, perhaps because the environment is less 'mutagenic' for sea-living mammals, or there are problems with the fossil record and the radiation of the modern whales occurred much more recently than is currently thought. A case for a less 'mutagenic' environment in the sea could be made on the basis of levels of naturally occurring radioactivity. Pelagic sea animals are not exposed to the largest source of irradiation for land organisms, which is the soil (derived from granite rocks). The contribution from cosmic rays is probably also smaller for sea animals, as their effect is already reduced by 70% at 10 m below sea level¹³. On the other hand, there is a higher contribution from both internalized ⁴⁰K and ²¹⁰Po, the effects of which are difficult to judge. ²¹⁰Po emits α -particles and can therefore deliver high doses of radiation in those tissues where it is accumulated, though it is not known whether it could accumulate in the gonads of the whales¹³.

Turning to the fossil record, the oldest fossils that are clear predecessors of the odontocete and mysticete whales come from

the early Miocene and might be as young as 20 Myr. Older fossils cannot be unequivocally ascribed to direct ancestors and could, therefore, represent extinct parallel lineages¹⁰. Given this, the fossil record could be compatible with an explosive radiation of all modern whale families around the early Miocene. This would bring their age down to about 20–25 Myr and the divergence rate for neutral nucleotide positions to about 0.15% Myr⁻¹, which is closer to the primate rate. Furthermore, our data support the notion that the major odontocete families and the mysticetes evolved at around the same time, as the average divergence among the odontocete families is similar to that between the odontocetes and mysticetes (Fig. 3). A picture thus emerges suggesting that the major families of modern whales may have evolved at about the same time during the early Miocene, although it will be necessary to support this inference by further data.

Our findings may also be of considerable practical value. Simple-sequence DNA fingerprints have several potential advantages over minisatellite fingerprints, particularly for population genetic work¹⁴. The main problem with SSLP analysis is that a suitable set of primers has to be developed for each species, which can be laborious and time consuming. Our results show that this is apparently not necessary for whales. Systematic monitoring of whale populations could therefore benefit greatly from the use of this method, in particular as only non-destructive sampling methods¹⁵ could be used. □

Received 29 July; accepted 1 October 1991.

1. Tautz, D. & Renz, M. *Nucleic Acids Res.* **12**, 4127–4138 (1984).
2. Tautz, D. *Nucleic Acids Res.* **17**, 6463–6471 (1989).
3. Weber, J. L. & May, P. E. *Am. J. hum. Genet.* **44**, 388–396 (1989).
4. Litt, M. & Luty, J. A. *Am. J. hum. Genet.* **44**, 397–401 (1989).
5. Amos, W., Barrett, J. A. & Dover, G. A. *Rep. int. Whal. Commun. (special issue)* **13**, 255–268 (1991).
6. Moore, S. S. *et al. Genomics* **10**, 654–660 (1991).
7. Wilson, A. C., Ochman, H., Prager, E. M. *Trends Genet.* **3**, 241–247 (1987).
8. Li, W.-H., Taniguchi, M. & Sharp, P. M. *J. molec. Evol.* **25**, 330–342 (1987).
9. Bulmer, M. *Nature* **325**, 728–730 (1987).
10. Barnes, L. G., Domning, D. P., & Ray, C. E. *Marine Mammal Sci.* **1**(1), 15–53 (1985).
11. Fordyce, R. E. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **31**, 319–336 (1980).
12. Hoelzel, A. R. & Dover, G. *Rep. int. Whal. Commun. (special issue)* **13**, 171–181 (1991).

13. Pentreath, R. J., Woodhead, D. S., Harvey, B. R. & Ibbett, R. D. *Proc. 3rd NEA Seminar 'Marine Ecology'* (OECD/NEA, Paris, 1979).
14. Tautz, D. *BioEssays* **12**, 44–46 (1990).
15. Hoelzel, A. R. & Amos, W. *Nature* **333**, 305 (1988).
16. Sakai, R. K. *et al. Science* **239**, 487–491 (1988).
17. Tautz, D., Trick, M. & Dover, G. *Nature* **322**, 652–656 (1986).

ACKNOWLEDGEMENTS. Tissue samples and DNA samples for this study were donated by S. Wada, J. Sigurjohnsson and R. Hoelzel. Pilot whale samples were collected with the help of D. Bloch, G. Desportes, R. Mouritsen and many others in a programme of research funded by the Faeroese government. We also thank S. Pääbo for valuable comments on the manuscript and W. Feldt for advice on the naturally occurring radioactivity in sea water.

Cloning and expression of a functional serotonin transporter from rat brain

Randy D. Blakely^{*†}, Haley E. Berson[†],
Robert T. Freneau Jr[‡], Marc G. Caron[§],
Margaret M. Peek^{*}, Heather K. Prince[†]
& Chris C. Bradley[†]

^{*} Department of Anatomy and Cell Biology and

[†] Program in Neuroscience, Emory University School of Medicine,
Atlanta, Georgia 30322, USA

Departments of [‡] Neurobiology, [§] Cell Biology and Medicine,
Duke University School of Medicine, Durham, North Carolina 27710, USA

SELECTIVE antagonism of serotonin (5-hydroxytryptamine, 5HT) and noradrenaline transport by antidepressants is a key element in the 'amine' hypothesis of affective disorders¹. Uptake^{2,3} and/or transport sites^{4,5} of 5HT have been reported to be reduced in platelets of patients suffering from depression and in post-mortem brain samples of depressed patients⁶ and suicide victims⁷. To date there has been little molecular information available on the structure and regulation of 5HT transporters. Using the polymerase chain reaction⁸ with degenerate oligonucleotides⁹ derived from two highly conserved regions of the transporters for noradrenaline¹⁰ and γ -aminobutyric acid¹¹ (GABA), we have identified a large family of related gene products expressed in rodent brain. One of these products hybridizes to a single 3.7-kilobase RNA restricted to rat midbrain and brainstem, where it is highly enriched within the serotonergic raphe complex. Transfection with a single 2.3-kilobase brainstem complementary DNA clone is sufficient to confer expression of a Na⁺-dependent 5HT transporter upon non-neural cells, with transport selectively and potently antagonized by 5HT uptake-specific antidepressants, including paroxetine, citalopram and fluoxetine.

Polymerase chain reaction (PCR) products from amplification of rodent and human cDNA of a size (~700 base pairs (bp)) compatible with the human noradrenaline and rat GABA transporters (hNAT¹⁰ and rGAT1¹¹, respectively) were purified, subcloned and sequenced. Eight unique clones were identified by sequencing (data not shown). Pairwise sequence comparison showed that most clones were as similar to each other as they were to hNAT and rGAT1, with ~50–60% identity. But another group, comprising clones rTB2-5 and rMB6-25, were more closely related to hNAT, with 84% and 67% identity, respectively. Given the significant overlap in antagonist sensitivity among monoamine neurotransmitter transporters¹², we surmised that these two species could be partial clones encoding dopamine and 5HT transporters. We focused our attention on rMB-25, using neuroanatomic techniques to ascertain the size and distribution of its messenger RNA in the rodent brain, and thus the likely substrate for this transporter homologue; we then designed probes to isolate a full-length functional clone.

In situ hybridization¹³ of endogenous RNA expression in slide-mounted sections of adult rat brain (Fig. 1a–c) reveals a prominent and specific hybridization signal to radiolabelled antisense cRNA transcribed from rMB6-25 overlying dorsal and median subdivisions of the serotonergic midbrain raphe complex¹⁴. Northern hybridization also indicates the presence of a single 3.7-kilobase RNA in rat midbrain and brainstem (Fig. 1d). Our inability to detect hybridization from total brain RNA underscores the restriction of gene expression for this putative transporter to cells of the mesencephalic and metencephalic raphe complex. Thus, the distribution in the central nervous system of hybridizing RNAs strongly suggests that rMB6-25 encodes a partial clone of the 5HT transporter. The adrenal RNA visualized in northern analyses is unlikely to arise

from cross-hybridization to the noradrenaline carrier as no RNAs were detected from the pheochromocytoma cells (PC12) derived from adrenal chromaffin cells or from human SK-N-SH neuroblastoma cells (Fig. 1d), both of which express high levels of the noradrenaline transporter¹⁰. In this regard, 5HT has been detected in mast cells lining rat adrenal arterioles¹⁵ and in a population of medullary cells synthesizing adrenaline^{16,17}. The properties of the adrenal RNA (equivalent size, detected using high-stringency hybridization) suggest that 3.7-kb mRNAs with high sequence correspondence to PCR clone rMB6-25 encode both brain and peripheral 5HT transporters.

A synthetic antisense oligonucleotide corresponding to the poorly conserved amino-acid sequence TIMAIFG (single-letter code) in the 5' end of PCR clone rMB6-25 was used to screen a rat brainstem cDNA library by plaque hybridization¹⁸. One positive plaque (BS4E-10) was identified from a total screen of 1.2×10^6 plaques, and the *Eco*RI insert subcloned into pBlue-script SKII⁺ (Stratagene). Figure 2a shows that HeLa fibroblasts transfected with the BS4E-10 cDNA, but not with the plasmid vector alone, express Na⁺-dependent 5HT uptake. 5HT transport was found to be saturable with substrate (data not shown), with an apparent Michaelis constant, K_m , of 320 nM. Transport assays in the presence of various uptake antagonists and substrates demonstrate a marked sensitivity of induced 5HT transport to tricyclic and heterocyclic antidepressants (Fig. 2b). The tertiary amine tricyclic antidepressants amitriptyline and imipramine were significantly more potent as antagonists (K_i = 16.9 and 18.7 nM, respectively) than their respective secondary amine congeners, nortriptyline and desipramine (K_i = 73.5 and 567 nM, respectively), giving a rank order potency of amitriptyline > imipramine > nortriptyline > desipramine. By contrast, the cloned noradrenaline transporter exhibits a reverse rank order potency of desipramine > nortriptyline > imipramine > amitriptyline¹⁰. The order and magnitude of tricyclic potency are generally equivalent to those obtained from serotonin transport studies in brain preparations¹⁹. As with brain preparations, halogenation of imipramine to chlorimipramine increases potency for antagonism in transfected cells by more than fivefold. Several of the nontricyclic antidepressants are considerably more selective for inhibition of endogenous 5HT over catecholamine transport, including fluoxetine²⁰, citalopram²¹ and paroxetine²². In this regard, paroxetine has a K_i of 0.39 nM for inhibition of 5HT transport after transfection with BS4E-10, nearly three orders of magnitude more potent than its inhibition of the cloned noradrenaline carrier¹⁰. The nonselective monoamine transport antagonist cocaine²³ also blocked 5HT transport induced by the cloned cDNA. The selective dopamine and noradrenaline transport inhibitors, GBR12909 and mazindol, only weakly inhibited 5HT uptake (K_i = 3.9 and 10.0 μ M, respectively). Thus, the activity of the protein encoded by the cloned cDNA, which we term rSERT, is pharmacologically very similar to the rat brain 5HT transporter, with high-affinity sites for both tricyclic and (the more selective) heterocyclic antidepressant antagonists.

The sequence of BS4E-10 reveals an open reading frame of 1,821-bp within a 2,278-bp cDNA (Fig. 3). The first ATG in the cDNA was assigned as the initiation codon on the basis of the initiation consensus sequence of Kozak²⁴. The open reading frame predicts a protein of 607 amino acids with a relative molecular mass of 68,000 (M_r 68K) and is distinguished by the presence of 11 or 12 regions of significantly extended hydrophobicity suitable for the formation of transmembrane domains²⁵. Two canonical sites for N-linked glycosylation are present on a large hydrophilic domain between putative transmembrane domains 3 and 4, which is similar to the location of a predicted extracellular loop in the cloned noradrenaline¹⁰ and GABA^{11,26} transporters. As with these carriers, the amino terminus fails to score as a signal sequence for membrane insertion²⁷ suggesting that it is retained in the cytoplasm. One consensus site for cyclic AMP-dependent protein kinase phosphorylation²⁸ is present

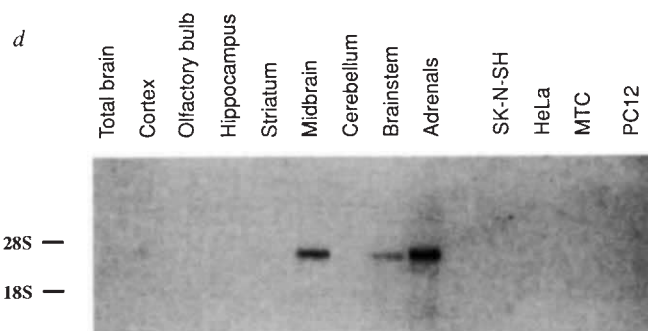
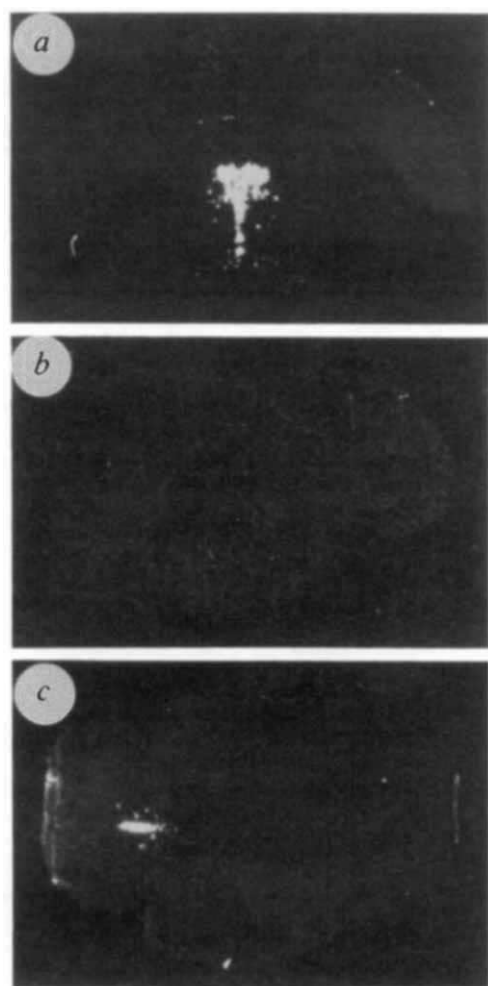


FIG. 1 *a-c*, *In situ* localization of RNA hybridizing to PCR clone rMB6-25 in rat coronal (*a* and *b*) and horizontal (*c*) brain sections. Sections presented in *a* and *c* were hybridized with ^{35}S -labelled antisense cRNA probe. Coronal section in *b* was taken adjacent to section displayed in *a* but probed with ^{35}S -labelled sense cRNA probe. Intense hybridization is confined to cells of the median and dorsal raphe nuclei. *d*, Distribution of RNA hybridizing to PCR clone rMB6-25. A 3.7-kb RNA is identified in rat midbrain, brainstem and adrenal gland. Human SK-N-SH cells and rat pheochromocytoma (PC-12) cells express the noradrenaline transporter¹⁸; human HeLa fibroblasts and rat CA-66 medullary thyroid carcinoma cells (MTC) serve as negative cell line controls.

METHODS. For *in situ* hybridization experiments, synthetic ^{35}S -labelled cRNA was synthesized with PCR fragment rMB6-25, cloned into the *Xba*I and *Pst*I sites of pBluescript SKII⁺, from either the T3 promoter after plasmid linearization with *Xba*I (antisense cRNA) or from the T7 promoter after linearization with *Xba*I (sense cRNA). cRNA synthesis and *in situ* hybridization to 4% paraformaldehyde-fixed rat brain sections was as previously described¹³. cDNA derived from PCR fragment rMB6-25 (100 ng) was radio-labelled with ^{32}P -labelled dCTP (50 μCi) using random oligonucleotide primers and hybridized to a nylon (Zetaprobe, BioRad) transfer of total RNAs (20 μg) derived from rat tissues and rat and human cell lines. Blots were prehybridized at 42 °C in 50% formamide, 5 $\times$ SSPE, 5 $\times$ Denhardt's solution, 10% dextran sulphate, 1% SDS and 100 $\mu\text{g ml}^{-1}$ salmon sperm DNA for 2 h, the probe was added and hybridization continued for 14 h. Blots were rinsed with two, 20-min washes at 22 °C in 2 $\times$ SSPE and 0.1% SDS, followed by a 1-h rinse at 65 °C in 0.1 $\times$ SSPE and 0.1% SDS, and then exposed to autoradiographic film with intensifying screen for 5 days. Positions of 18S (1,950 kb) and 28S (4,700 kb) ribosomal RNA markers are indicated. All lanes carry equivalent loads on the basis of even intensity of rRNAs.

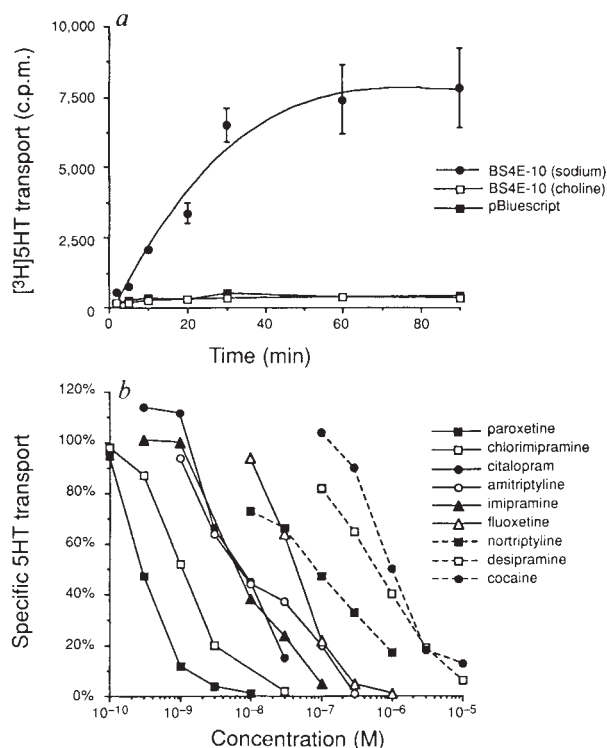


FIG. 2 *a*, Presence of sodium-dependent 5HT transport in HeLa fibroblasts transfected with BS4E-10 cDNA. *b*, Inhibition of 5HT transport in transfected cells by antagonists of monoamine transport.

METHODS. The cDNA (BS4E-10) insert was excised from $\lambda\text{gt}10$ with *Eco*RI and subcloned in pBluescript SKII⁺ (Stratagene) placing the presumptive amino terminus (determined from PCR amplification and sequencing) immediately downstream of the T7 RNA polymerase promoter. Cells (10^5 per well) were infected with recombinant vaccinia virus strain VTF7-3 (ref. 38) expressing T7 RNA polymerase³⁹, followed 30 min later by liposome-mediated (Lipofectin, BRL) transfection of the cDNA construct. Control transfections consisted of equivalent amounts of transfected vector alone. 5HT transport was assayed 8 h after transfection as described for the analysis of the transfected noradrenaline carrier¹⁰, using 5-[1,2- $^3\text{H}(\text{N})$]hydroxytryptamine creatine sulphate ($[^3\text{H}]5\text{HT}$, at 20 nM; Dupont/NEN) as substrate in Krebs/Ringers/Tris/HEPES (KRTH) uptake medium³⁹. Assays were terminated and washed with cold KRTH, cells were solubilized with 1% SDS and the accumulated radioactivity determined by scintillation counting. Data are presented as c.p.m. per well and represent mean $\pm$ s.e.m. of triplicate experiments. Sodium-dependence was determined by isotonic substitution of assay NaCl with choline chloride. Inhibition assays in duplicate or triplicate were run with or without increasing concentrations of selected substrates and antagonists of 5HT, noradrenaline and dopamine transport. Nonspecific transport was assessed with a parallel transfection of pBluescript for each assay and values subtracted from signals obtained with BS4E-10 cDNA. Inhibition data are presented as a percentage of 5HT uptake obtained with labelled substrate alone. Errors associated with independent experiments were less than 10% of the mean values plotted. Dopamine, noradrenaline, adrenaline, and histamine were ineffective in blocking 5HT transport induced by BS4E-10 ($K_i > 10 \mu\text{M}$). Experiments with increasing concentrations of unlabelled 5HT yielded a K_m of 320 nM and a V_{max} of 3.1×10^{-18} mol per cell per min.

FIG. 3 Nucleotide and deduced amino-acid sequence (single-letter code) of the rat 5HT transporter (rSERT) encoded by BS4E-10. Nucleotides are numbered in the 5' to 3' direction beginning with the first nucleotide in the cDNA clone. The presumptive start of translation begins at nucleotide 48, with the first in-frame stop codon at nucleotide 1,869, yielding an 1,821-bp open reading frame encoding a 607-amino-acid polypeptide. A second in-frame ATG is present at base 111, but surrounding sequences present a poorer consensus sequence for initiation; initiation from the upstream ATG also reveals several absolutely conserved residues with the noradrenaline carrier within the more extended amino terminus. Twelve stretches of amino-acid sequence predicted to encode transmembrane domains²⁵ are underlined. One canonical cAMP-dependent protein kinase recognition site²⁸ and two canonical sites for *N*-linked glycosylation are shaded and double-underlined, respectively. Sequences from bases 279–974 match those obtained from the partial cDNA clone rMB6-25.

METHODS. The conserved amino-acid sequences NVWRFPY and WIDAAATQ, of hNAT (ref. 10) and rGAT1 (ref. 11), were used to design degenerate inosine (I)-substituted oligonucleotides of sequence 5'-CCGCTCGAGAA(C/T)GT(G/C)TT(C/T)CC-(A/G/C/T)TA-3' and 5'-GCTCTAGAGCTG-(A/G)GTIGC(A/G)GC(A/G)TC(A/G)A(T/G)CC-A-3', respectively. (Underlined sequence indicates addition of 5' restriction sites for cloning.) Oligonucleotides were combined with single-stranded rat and human cDNAs, synthesized from poly(A)⁺RNA with random hexamer primers (Amersham) in PCR reactions conducted with *Taq* polymerase (Promega) for 30 cycles at 94 °C for 1 min, 45 °C for 2 min, 72 °C for 3 min, with the final extension lengthened to 15 min. Products of ~700 bp isolated by phenol extraction and ethanol precipitation were digested with *Eco*RI to prevent recloning of rGAT1, which has an *Eco*RI site between the oligonucleotides used for amplification, and digested with *Xba*I and *Xho*I to produce staggered cloning ends. Samples were gel-purified (GENECLEAN, Bio101), and ligated into *Xba*I-*Xho*I-digested pBluescript SKII⁺ (Stratagene). Partial sequencing of double-stranded plasmid clones was achieved by dideoxynucleotide chain termination using Sequenase (US Biochem). Alignment of cloned fragments to hNAT and rGAT1 was achieved iteratively with the program BESTFIT of the Wisconsin GCG software package³⁰. Using an end-labelled oligonucleotide derived from a non-conserved region of the 5' end of PCR clone rMB6-25, we isolated a single positive plaque in a screen of 1.2×10^6 bacteriophage from a rat brainstem cDNA library prepared in λ gt10 (Clontech). The insert was liberated from purified bacteriophage with *Eco*RI and ligated into *Eco*RI-digested pBluescript SKII⁺. Dideoxynucleotide chain termination sequencing was achieved on both strands with Sequenase (US Biochem). Sequences obtained from two separate partial cDNAs isolated from a rat midbrain cDNA library in a separate screen were also used to confirm the sequence and interpret compressions. MacVector DNA analysis software (IBI) was used for sequence assembly and analysis.

```

      20      40      60      80      100
TCAGAAAGTCTGTAGAGTGAAGGACAGAGGAGCTGTCAAGAAATGGTGTCTACAGAGGTTGCCACACAGCGGACAGGCAGAGCCTAGC
      *      *      *      *      *
      120      140      160      180      200
CAAATATCCAATGGGTACTCTGCAGTCCCCAGGCACAAGTGCAGGGGAGCAAGCTTCACACTCGATCCAGCTGCCACCAACCCCTGGTGGCTGAGATT
      *      *      *      *      *
      220      240      260      280      300
CGCCAAAGGGAGCGGGAGACCTGGGGCAAGATGGATTCTCTCTGTCGCTCATTGGCTATGCCGTGGACCTGGGCAACATCTGGCGGTTCTCTTACA
      *      *      *      *      *
      320      340      360      380      400
TATGCTACCAAGATGGCGAGGGCCCTCTCTCTCTATACCATATGGCCATTTTCGGGGGATCCCGCTCTTTTACATGGAGCTCGCACTGGGCCA
      *      *      *      *      *
      420      440      460      480      500
GTACCACCGAAACGGGTGCATTTCATATGGAGGAAGATCTGCCGATTTTCAAAGGCATTGGTTACGCCATCTGCATCATCGCCTTTTACATCGCCTCC
      *      *      *      *      *
      520      540      560      580      600
TACTACAACACCATCATAGCTGGGCGCTCTACTACCTCATCTCTCTCTCAGGACCGGCTGCCCTGGACAGCTGCACGAACCTCTGGAACACTGGCA
      *      *      *      *      *
      620      640      660      680      700
ACTGCACCAACTACTTCCGCCAGGACAAACATCACTGGAGCTGCATTCACGTCCTCCGCTGAGGAGTTCTACTTGGCCATGTCTGCAGATCCACCA
      *      *      *      *      *
      720      740      760      780      800
GTCTAAGGAGCTCCAGGACCTGGGACCATCAGCTGGAGCTGACTCTCTGCATCGTCTCATCTTACCCTAATCTACTTTAGCATCTGGAAGGGGCTC
      *      *      *      *      *
      820      840      860      880      900
AAAACATCTGGCAAGGTGGTGTGGGTGACAGCCACCTTCCCATACATGTCTCTCTGCTGCTGGTGGAGGGGGCCACCTTCTGGAGCCTGGAGAG
      *      *      *      *      *
      920      940      960      980      1000
GGGTCTGCTTCTACTTGAACCAACTGGCAGAACTCTTGGAGACAGGGGTGGGTAGATGCCCGCTCAGATCTCTTCTCTTGGCCCGGGCTT
      *      *      *      *      *
      1020      1040      1060      1080      1100
TGGGGTCTCTGGCTTTTGGTGTAGTACAACAAGTTCACAACAAGTGTACCAAGATGCCCTGGTACCACTGTGGTGAAGTGCATGACAAAGCTTCGTC
      *      *      *      *      *
      1120      1140      1160      1180      1200
TCTGGCTTCGTCATCTTACCGGTGCTTGGCTACATGGCGGAGATGAGGAATGAAGATGTGTACAGAGTGGCCAAAGACCGAGGCCACGCTCTCTTCA
      *      *      *      *      *
      1220      1240      1260      1280      1300
TCACGTATGCAGAGGCAATAGCAACATGCCAGCATCCAGCTTCTTGGCATCATCTTCTCTCATGTTAATACGCTGGGATGGAGACGACGTTCCG
      *      *      *      *      *
      1320      1340      1360      1380      1400
AGGCCTGGAAGGTGTATGATCAGCTGTGCTGATGAGTTCCTCACATCTGGGCAAGCGCAGGGAATGGTTCGTGCTCATCGTGGTATCAGTGGCTG
      *      *      *      *      *
      1420      1440      1460      1480      1500
TTGGGATCCCTGCTCACACTGACGTGAGGAGGGGACGTGAGTGTGCTGGAGGAGTATGCCAGGGGCCAGCAGTGTACCGTGGCCCTCATCG
      *      *      *      *      *
      1520      1540      1560      1580      1600
AGGCCGTCGCGGTGCTTGGTCTTATGGAATCACTAGTCTGTCAGCGATGTGAAGGAGATGCTGGGCTTCAGCCCGGATGGTTTGGAGGATCTGCTG
      *      *      *      *      *
      1620      1640      1660      1680      1700
GGTGGCCATCAGCCCTCTGTTCTCTCTGTCATCATTTGCAGTCTTCTGATGAGCCACCCAGCTACGGCTTTTCCAATACAATATCCCACTGGAGT
      *      *      *      *      *
      1720      1740      1760      1780      1800
ATCGTCTTGGGCTACTGCATAGGATGTGCTCCGTCATCTGCATCCCTACCTATATCATTTATCGGTGATCAGCACTCCGGGACACTTAAGGAGCGCA
      *      *      *      *      *
      1820      1840      1860      1880      1900
TTATTAAGATATCACTCTGAAACACCCACAGAAATCCCGTGGGGACATCCGCATGAATGCTGTGTAACACACCTGGGAGAGGACACCTCTTCCCA
      *      *      *      *      *
      1920      1940      1960      1980      2000
GCCACCTCTCTCAGCTCTGAAAGCCCACTGGACTCTCCCTCTAAGCCAAAGCCTGATGAAGACAGGCTCTAACCCTATGGTGCCAGACTCTTGT
      *      *      *      *      *
      2020      2040      2060      2080      2100
GGATTCCGACCACTTCTTCCGTGGACTCTCAGACATGCTACCACTTCGATGGTGACACCACTGAGCTGGCTCTTGGACACCTCAGGAGTGAAGGA
      *      *      *      *      *
      2120      2140      2160      2180      2200
GGGATGAACGCCACCACTCATCAGCTAGCTTTCAGGTTAGAATTAGGTCTGTGAGAGTCTGTATCATGTTTGGTAAAGATCATATACCCCGCATCTG
      *      *      *      *      *
      2220
TTAGCTCTAAAGCCTTCAATGTTTCATGAATACATAAACCACTAAGAGAAACAGAGATGCTTCTGCTAGCCATATAT

```


a

rSERT	1	MVFYRRVSPQRTGQSLAKYPMGTLSQPGT..SAGDEASHSIPAATTTLV	48
hNAT	1	MLLARMPNQVQFENNADTGPQPLRARKTAELLVVKERNVGQCLLAPR.	49
rGAT1	1	..MATDNKSVADGQISTEVS.EAPVADK.PKTLVVKVQK.....KA	38
rSERT	49	AEIRQGERETWKKMDPLLSVIGYAVDLGNIRFPYICYNCCGAFLLPY	98
hNAT	50	.DGDAPRETWGKKIDPLLSVVGAVDLANVRFPYLCYKNGGAFLLPY	98
rGAT1	39	..GDLPRDRTWKGRDFLMSCVGYAIGLGNVRRFPYLCCKNGGAFLLPY	86
rSERT	99	TIMALFGGIFLPMELALGQYHRNCISIRKICIFKGIYAIICIAFY	148
hNAT	99	TLFLTIAGMPLFYMELALGQYNREGAATVM.KICEFFKGVGYAVILIALY	147
rGAT1	87	FLTLTFAGVPELLECSLGOYTSIGGLGVN.KLAPMKGVLSAAAVLSFW	135
rSERT	149	IASYNTITLAWLYLISLTDRLPWTSCNTSNNTGNCI....NYFAQ.	192
hNAT	148	VGFIYNVILIAWSLYLFSFTLNLNPTDCGHTWNSPNCIDPKLLNGSVLG	197
rGAT1	136	LNIIYIVILISWAIYYLYNSFTTTLPWKQCDNPNNTDRCS....NYSLVN	181
rSERT	193	DNITWTLHSTPAEEFYLRHVLQIHQSKGLQDLGTISWOLTLICVLIPTV	242
hNAT	198	NHTKYSKYKFTPAEEFYERGVHLHHESSGIHDIQLPQWOLLCLMVVIV	247
rGAT1	182	TNNMTS....AVVEFWERNMHQMTD..GLDKPGQIRNPLAITLAIWVL	224
rSERT	243	LYESLWKGVKTSKGKVVVWITATFPYIVLSVLLVRGATLPGAWRGVVFYLP	292
hNAT	248	LYESLWKGVKTSKGKVVVWITATFPYIVLVLLVHGVTLPASNGINAYLHI	297
rGAT1	225	VVEFCIWKGVGTGKVVVYSATYPIMLIILFRGVTLPGAKEGILFYITP	274
rSERT	293	NWQKLETCGVVVDAAQIFPSSLGPGGVLLAFASYNKFNNNCYQDALVTS	342
hNAT	298	DFYREKEATVWIDAAQTQIFPSSLGAGFGVLLAFASYNKFDNNCYRDALLTS	347
rGAT1	275	NFRKLSDSVEVWLDAAQTQIFPSSYGLGLSLIALGSYNSFHNIVYRDSIIVC	324
rSERT	343	VVNCMTSFVSGEVIPTVLGYMAEMRNEVDSEVAKDAGPSLLEITVAAIA	392
hNAT	348	SINCITSFVSGFAIFSLILGYMAHEHKVNIEDVATE.CAGLVEILYPEAIS	396
rGAT1	325	CINSCISMFAGEVIFSIIVGMAHVTKRSIADVAAS..GPGLAFLAYPEAVT	373
rSERT	393	NMPASTFFAIIFFILMITLGLDSTFAGLEGVITAVLDEFPHIWAQRREWF	442
hNAT	397	TLSCGTFFWAVVEFVMLLALGLDSSMGMEAVITGLADDF.QVLKRHRKLF	445
rGAT1	374	QLPISPLWAILFFSMILMLGLDSSQFCTVEGFTALVDEYPRLLNRNREL	423
rSERT	443	VLIVVITCVLGSLLTLTSGGAVVVTILEEYAT.GPAVLTVALIEAVAVSW	491
hNAT	446	TFGVTFSTFLALFCITKGGIIVLTLDTFAA.GTSLFAVLMEALGVSW	494
rGAT1	424	IAAVCSYSLIGISNITOGGIIVFKLFDYSSASGMSLLFLVFEECVSISN	473
rSERT	492	FYGITQPCSDVKEMLQFSFGWFRICWVAISLFLFFIICSFLMSPPOQR	541
hNAT	495	FYGVDRFESNDIQMGFRGLYWRICWKEFVSFAFLFVVVSIINFKPLT	544
rGAT1	474	FYGVNRFYDNIQMGFRICWVAISLFLFFIICSFLMSPPOQR	523
rSERT	542	LFQYNYPHHSIVLGYCIGMSS.VICITYIIRLISTPGLTKERIKSIT	590
hNAT	545	YDDYIFPPHANWVGNGIALSEMVL.VPIYVIKFLSTQGSLSLWRLAYGCT	593
rGAT1	524	MGSYVFPKVGQGVGLMALSSMVL.IPGYMAFMFLTLKGLSKQLQVMIQ	572
rSERT	591	PETPTEIPCG.DIRMANV.....	607
hNAT	594	PENEHHLVAQRDIRQQLHLWLAI.....	611
rGAT1	573	P.SEDIVRPENGPEQPGSSASKAYI	599

b

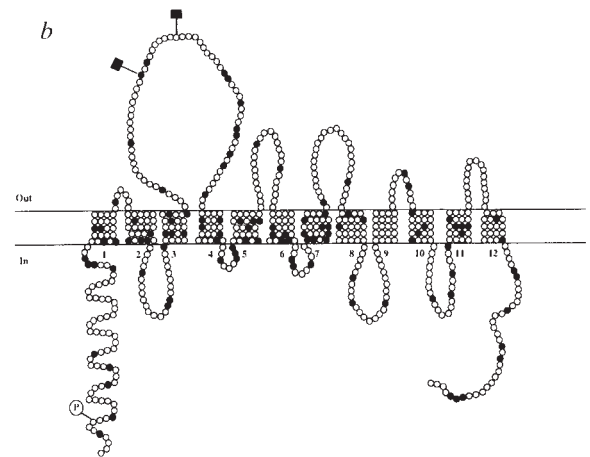


FIG. 4 Alignment of amino-acid sequences encoding rat 5HT (rSERT), human noradrenaline (hNAT), and rat GABA (rGAT1) transporters. *a*, Amino-acid sequence alignments were produced by iterative use of the BESTFIT routine of the Wisconsin GCG software package³⁰. Shaded regions reflect sequences absolutely conserved among all three carriers, representing ~30% of amino-acid residues. Solid lines above rSERT reflects location of putative transmembrane domains. Note the high degree of conservation among the three carriers between amino acids 76 and 98, where 17/23 residues are conserved. *b*, Structural model for the 5HT transporter. Circles represent individual amino acids. Shaded circles reflect amino acids absolutely conserved between rSERT and hNAT, but not present in the related rGAT1, which are possibly involved in transport activities specific to noradrenaline and serotonin carriers like the binding of tricyclic antidepressants. Note the localization of conserved residues in certain transmembrane domains (numbered 1–12) relative to predicted cytoplasmic and extracellular domains. Black boxes indicate sugars attached to canonical sites for N-linked glycosylation. A circled P indicates a canonical site (RRXS) for cAMP-dependent protein kinase recognition.

near the end of the N terminus. Interestingly, 5HT transporters derived from a human placental choriocarcinoma cell line (JAR) are regulated by cAMP²⁹.

Comparison³⁰ of the predicted amino-acid sequences of rSERT, hNAT and rGAT1 demonstrates a striking conservation of sequence (Fig. 4a). Although rSERT is more closely related to hNAT than to rGAT1, with 50% (versus 43%) of residues being absolutely conserved, rising to 71% (versus 67%) similarity when conservative substitutions are accepted, about 30% of all residues are absolutely conserved across the three carriers. Many of these residues are positioned in or adjacent to the transmembrane domains and probably help create the secondary structure necessary for ion binding and/or substrate translocation. To determine which amino acids could be involved specifically in monoamine transport, for example those that bind tricyclic antidepressants and cocaine, we established the positions of residues that are absolutely conserved in rSERT and hNAT but are not conserved in rGAT1, because the latter transporter is insensitive to these agents. Superimposed on a preliminary structural model of rSERT (Fig. 4b), these residues cluster prominently in several putative transmembrane domains, particularly transmembrane domains 5–7 (compare with domains 9 and 12). Interestingly, there is only one acidic residue (Asp 75 in transmembrane domain 1) in the transmembrane domains that is conserved between rSERT and hNAT but is absent from rGAT1. Most transport antagonists are believed to occupy sites overlapping the substrate-binding site^{31–33}. A negatively charged residue may be involved in the binding of polar

amino groups of substrates and antagonists to monoamine transporters^{19,34}, so we suggest that monoamine neurotransmitter transporters bind their substrates and antagonists in the plane of the membrane, possibly by means of determinants of the transmembrane domains 1 and 5–7. A similar extended intramembrane pocket has been proposed for binding monoamine neurotransmitters and antagonists to G protein coupled receptors^{35,36}. Apart from the similarities with the noradrenaline and GABA transporters, no other significant identity was found in sequence comparisons using the GenBank data base, which included receptors, the Na⁺/glucose and Na⁺/proline transporters, and facilitated carriers.

In summary, we have identified a single cDNA from brain which is sufficient to form a fully functional 5HT transporter in nonneuronal cells. Higher resolution of the spatial organization of important residues in the 5HT transporter should assist in the design of more selective therapeutic agents, and the high-affinity tricyclic-antidepressant binding sites on rSERT and hNAT should help to define key residues in antagonist selectivity. Selective serotonin transport inhibitors are now being prescribed for the clinical management of depression, obsessive-compulsive disorder, panic disorder, bulimia and obesity³⁷. The cloning of rSERT will enable transcriptional and post-translational regulation of the 5HT transporter to be studied directly in animal models and could lead to the identification of a human homologue that could be used to assess genetic disturbances underlying neuropsychiatric disorders.

Received 23 August; accepted 9 October 1991.

1. Ashton, H. *Brain Systems, Disorders and Psychotropic Drugs* 283–330 (Oxford University Press, New York, 1987).
2. Meltzer, H. Y., Arora, R. C., Baber, R. & Tricou, B. J. *Arch. gen. Psychiat.* **38**, 1322–1326 (1981).
3. Suranyi-Cadotte, B. E., Quiron, R., Nair, N. P. V., Lafaille, F. & Schwartz, G. *Life Sci.* **36**, 795–799 (1985).
4. Briley, M. S., Langer, S. Z., Raisman, R., Sechter, D. & Zarifan, E. *Science* **209**, 303–305 (1980).
5. Paul, S. M., Rehavi, M., Skolnick, P., Ballenger, J. C. & Goodwin, F. K. *Arch. gen. Psychiat.* **38**, 1315–1317 (1981).
6. Perry, E. K., Marshall, E. F., Blessed, G., Tomlinson, B. E. & Perry, R. H. *Brit. J. Psychiat.* **142**, 188–192 (1983).
7. Stanley, M., Virgilio, J. & Gershon, S. *Science* **216**, 1337–1339 (1982).
8. Saiki, R. K. *et al. Science* **239**, 487–494 (1988).
9. Rathe, R. J. *molec. Biol.* **183**, 1–12 (1985).
10. Pacholczyk, T., Blakely, R. D. & Amara, S. G. *Nature* **350**, 350–354 (1991).
11. Guastella, J. *et al. Science* **249**, 1303–1306 (1990).
12. Richelson, E. *Mayo Clin. Proc.* **65**, 1227–1236 (1990).
13. Freneau, R. T. Jr *et al. Proc. natn. Acad. Sci. U.S.A.* **88**, 3772–3776 (1991).
14. Steinbusch, H. W. M. & Nieuwenhuis, R. in *Chemical Neuroanatomy* (ed. Emson, P. C.) 131–207 (Raven, New York, 1983).
15. Hinson, J. P., Vinson, G. P., Pudney, J. & Whitehouse, B. J. *J. Endocr.* **121**, 253–260 (1989).
16. Verhofstad, A. A. J. & Jonsson, G. *Neuroscience* **10**, 1443–1453 (1983).
17. Holworth, M. A. & Brownfield, M. S. *Neuroendocrinology* **41**, 230–236 (1985).
18. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* 2nd edn (Cold Spring Harbor Laboratory, New York, 1989).
19. Maxwell, R. A. & White, H. L. in *Handbook of Psychopharmacology* (eds Iversen, L. L., Iversen, S. D. & Snyder, S. H.) 83–155 (Plenum, New York, 1978).

20. Lemberger, L. H. *et al. Clin. Pharmac. Ther.* **23**, 421–429 (1978).
21. Hyttel, J. *Psychopharmacology* **51**, 225–233 (1977).
22. Buss Lassen, J. *Eur. J. Pharmac.* **47**, 351–358 (1978).
23. Ritz, M. C., Cone, E. J. & Kuhar, M. J. *Life Sci.* **46**, 635–645 (1990).
24. Kozak, M. *Nucleic Acids Res.* **15**, 8125–8148 (1987).
25. Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105–132 (1982).
26. Nelson, H., Mandiyan, S. & Nelson, N. *FEBS Lett.* **269**, 181–184 (1990).
27. von Heijne, G. *Eur. J. Biochem.* **133**, 17–21 (1983).
28. Kemp, B. E. & Pearson, R. B. *Trends biochem. Sci.* **15**, 342–346 (1990).
29. Cool, D. R., Leibach, F. H., Bhalla, V., Maheshi, V. & Ganapathy, V. *J. biol. Chem.* **266**, 15750–15755 (1991).
30. Devereux, J., Haeberli, P. & Smithies, O. *Nucleic Acids Res.* **12**, 387–395 (1984).
31. Andersen, P. H. *Eur. J. Pharmac.* **166**, 493–504 (1989).
32. Graham, D., Esnaud, H., Habert, E. & Langer, S. Z. *Biochem. Pharmac.* **38**, 3819–3826 (1989).
33. Marcusson, J. O., Andersson, A. & Bäckström, I. *Psychopharmacology* **99**, 17–21 (1989).
34. Koe, B. K. *J. Pharmac. exp. Ther.* **199**, 649–661 (1976).
35. Strader, C. D., Sigal, I. S. & Dixon, R. A. F. *FASEB J.* **3**, 1825–1832 (1989).
36. Kobilka, B. K. *et al. Science* **240**, 1310–1316 (1988).
37. Fuller, R. W. & Wong, D. T. *Ann. N.Y. Acad. Sci.* **600**, 69–80 (1990).
38. Furst, T. R., Niles, E. G., Studier, F. W. & Moss, B. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8122–8126 (1986).
39. Blakely, R. D., Clark, J. A., Pacholczyk, T. & Amara, S. G. *Analyt. Biochem.* **194**, 302–308 (1991).

ACKNOWLEDGEMENTS. We thank J. Kiltey and S. Amara for gifts of reagents. This work was partially supported by a Woodruff Neuroscience Investigator Award from The Emory University School of Medicine and the National Institute of Drug Abuse (R.D.B.), by the National Institutes of Health (M.G.C. and R.T.F.), and an Award from Glaxo Research Laboratories (R.T.F.). C.C.B. is a student in the Emory Medical Scientist Training Program.

Cloning of cDNA for the glutamate-binding subunit of an NMDA receptor complex

Keshava N. Kumar^{*†}, Nanda Tilakaratne^{*†},
Peter S. Johnson^{*}, Annette E. Allen^{*}
& Elias K. Michaelis^{*†§}

^{*} Department of Pharmacology and Toxicology, and the [†] Center for Biomedical Research, 2099 Constant Avenue, West Campus, University of Kansas, Lawrence, Kansas 66047, USA

THE amino acids L-glutamic and L-aspartic acids form the most widespread excitatory transmitter network in mammalian brain^{1,2}. The excitation produced by L-glutamic acid is important in the early development of the nervous system^{3,4}, synaptic plasticity and memory formation^{5,6}, seizures^{7,8} and neuronal degeneration^{9,10}. The receptors activated by L-glutamic acid are a target for therapeutic intervention in neurodegenerative diseases, brain ischaemia and epilepsy. There are two types of receptors for the excitatory amino acids, those that lead to the opening of cation-selective channels and those that activate phospholipase C (ref. 11). The receptors activating ion channels are NMDA (N-methyl-D-aspartate) and kainate/AMPA (α -amino-3-hydroxy-5-methyl-4-isoxazole propionate)-sensitive receptors. The complementary DNAs for the kainate/AMPA receptor^{12,13} and for the metabotropic receptor¹⁴ have been cloned. We report here on the isolation and characterization of a protein complex of four major proteins that represents an intact complex of the NMDA receptor ion channel and on the cloning of the cDNA for one of the subunits of this receptor complex, the glutamate-binding protein.

We reported previously¹⁵ that a set of partially purified glutamate-binding proteins formed ligand-activated ion channels when reconstituted into liposomes. L-Glutamate and NMDA, but not kainate or quisqualate, increased the rate of cation influx into liposomes reconstituted with the glutamate-binding proteins and the NMDA receptor antagonist 2-amino-5-phosphonopentanoic acid completely blocked glutamate activation of this reconstituted ion channel. Electrophoretic analysis of the reconstituted proteins revealed enrichment of four major bands¹⁵. We have modified the procedures for the solubilization and purification of these proteins and have purified a complex of four proteins that has all the ligand binding characteristics of an NMDA receptor.

The fraction eluted from an L-glutamate-treated matrix by introducing 5 mM NMDA into the elution buffer contains proteins with estimated relative molecular masses (M_r) of 70,000 (70K), 62K, 43K and 41K (doublet), and 36K and 31K (doublet) (Fig. 1). The proteins of M_r about 41–43K and 31–36K

ststituted proteins revealed enrichment of four major bands¹⁵. We have modified the procedures for the solubilization and purification of these proteins and have purified a complex of four proteins that has all the ligand binding characteristics of an NMDA receptor.

The fraction eluted from an L-glutamate-treated matrix by introducing 5 mM NMDA into the elution buffer contains proteins with estimated relative molecular masses (M_r) of 70,000 (70K), 62K, 43K and 41K (doublet), and 36K and 31K (doublet) (Fig. 1). The proteins of M_r about 41–43K and 31–36K

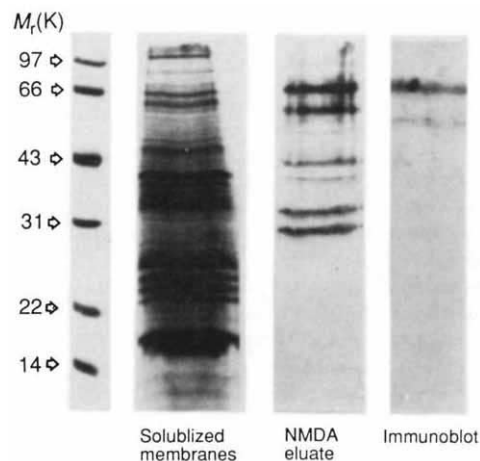


FIG. 1 Electrophoresis of the purified NMDA receptor complex and western blot analysis of the isolated proteins.

METHODS. Synaptic membranes were isolated from whole rat brain homogenates as described¹⁶; membranes were solubilized in 10 mM potassium phosphate buffer, pH 7.4, containing protease inhibitors (0.3 mM phenylmethane sulphonylfluoride, 10 mM ϵ -amino caproic acid, 0.1 mM EGTA, 0.1 mM benzamide), CHAPS (1.5% w/v), *n*-octylglucoside (0.5% w/v) and glycerol (10% v/v). The solubilized membrane proteins were loaded onto a ReactiGel column treated with L-glutamate. The column was washed with solubilization buffer and sequentially eluted with the same buffer containing first 30 mM KCl, then 5 mM NMDA, and finally 1 M KCl. Fractions eluted with 5 mM NMDA had the highest activity of NMDA-sensitive L-[³H]glutamate binding and of [³H]TCP and [³H]glycine binding. Synaptic membrane proteins (7 μ g protein), NMDA eluate (6 μ g protein), and molecular weight standards ($M_r \times 10^{-3}$; left-hand lane) were subjected to SDS-PAGE on 12% polyacrylamide gels. Gels were stained with silver nitrate. For western blotting, proteins were electrotransferred onto nitrocellulose and labelled with polyclonal antibodies raised against the previously purified 71K glutamate-binding protein¹⁷.

[‡] Present address: Department of Psychiatry, Medical College of Pennsylvania, EPPI, 3200 Henry Avenue, Philadelphia, Pennsylvania 19129, USA.

[§] To whom correspondence should be addressed.

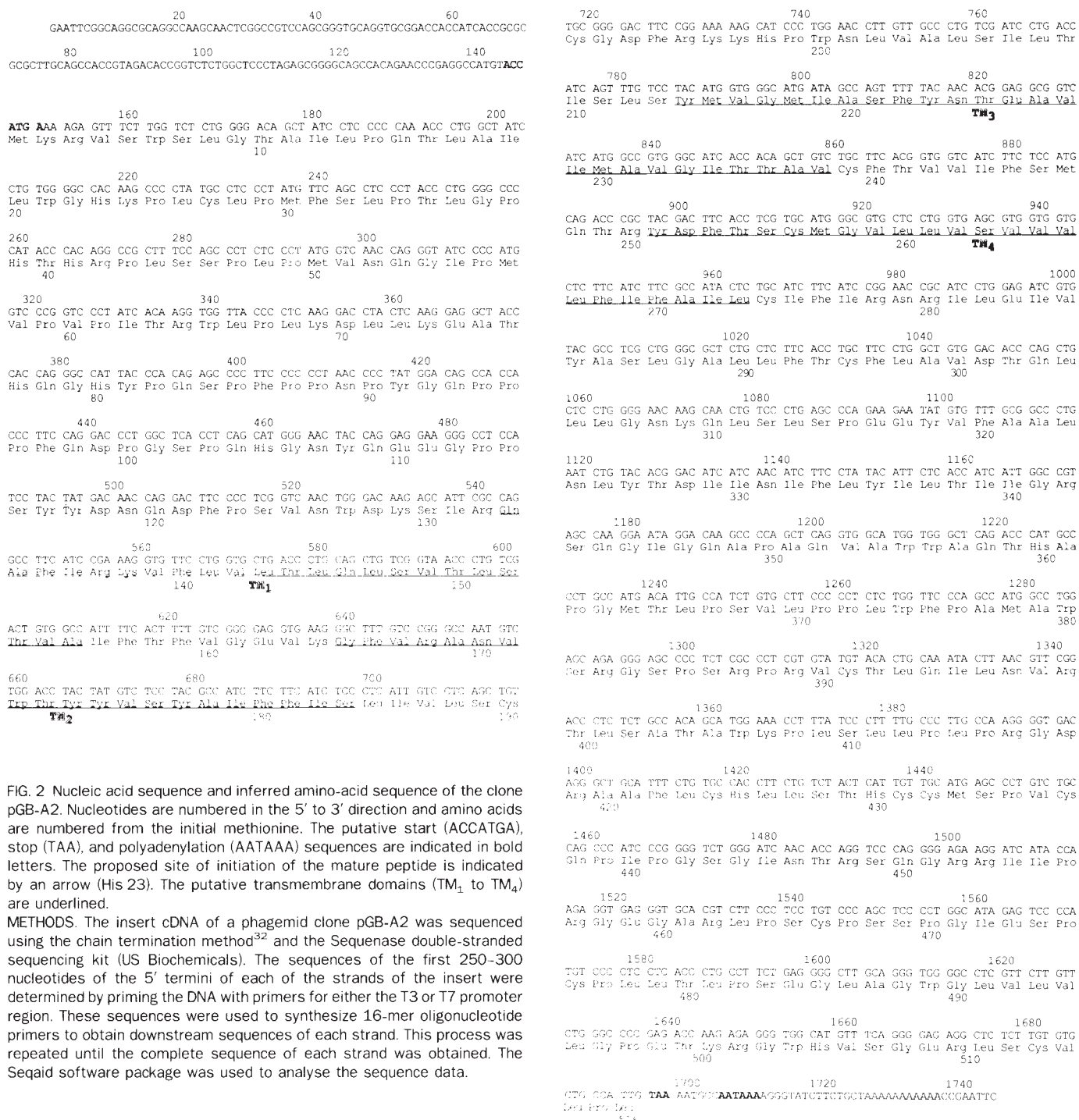


FIG. 2 Nucleic acid sequence and inferred amino-acid sequence of the clone pGB-A2. Nucleotides are numbered in the 5' to 3' direction and amino acids are numbered from the initial methionine. The putative start (ACCATGA), stop (TAA), and polyadenylation (AATAAA) sequences are indicated in bold letters. The proposed site of initiation of the mature peptide is indicated by an arrow (His 23). The putative transmembrane domains (TM₁ to TM₄) are underlined.

METHODS. The insert cDNA of a phagemid clone pGB-A2 was sequenced using the chain termination method³² and the Sequenase double-stranded sequencing kit (US Biochemicals). The sequences of the first 250–300 nucleotides of the 5' termini of each of the strands of the insert were determined by priming the DNA with primers for either the T3 or T7 promoter region. These sequences were used to synthesize 16-mer oligonucleotide primers to obtain downstream sequences of each strand. This process was repeated until the complete sequence of each strand was obtained. The Seqaid software package was used to analyse the sequence data.

frequently migrate in SDS-PAGE either as a diffuse band (see Fig. 6 of ref. 15) or as a doublet (Fig. 1). This complex of proteins represents >2,000-fold enrichment of NMDA-sensitive L-[³H]glutamate binding, of dizocilpine (MK-801)-sensitive [³H]TCP (*N*-(1-(2-thienyl)cyclohexyl)piperidine) binding, and of strychnine-insensitive [³H]glycine binding when compared with brain homogenates. Glutamate, glycine and polyamines accelerate the binding of [³H]TCP to the isolated complex (data not shown).

We previously raised specific antibodies against the 71K glutamate-binding protein (GBP) isolated from rat brain synaptic membranes^{16,17}. These antibodies react with the 68–70K protein in the isolated NMDA receptor complex (Fig. 1); they block glutamate activation of ion flux into liposomes¹⁵ and inhibit the neurotoxic effects of NMDA in primary cultures of brain hippo-

campal neurons¹⁸. We used both polyconal and monoclonal antibodies to screen a rat hippocampal cDNA expression library that was constructed by cloning the cDNAs into the *Eco*RI site of lambdaZap (Stratagene) phage vector. Twelve plaques were isolated as potential clones expressing this subunit of the receptor. Cells were extracted after activation of the lacZ promoter with isopropylthiogalactoside (IPTG) from colonies of *Escherichia coli* XL-1 blue strains potentially harbouring the phagemid for the GBP introduced by automatic excision¹⁹. Extracts that reacted positively in dot-blot immunostaining with monoclonal antibodies were analysed by western blotting. We selected one clone, pGB-A2, which contained an insert of ~1.8 kilobases (kb) and yielded an antibody-labelled peptide of about 60 K.

The 1,742-nucleotide sequence of the cloned cDNA is in frame

with that of β -galactosidase and is shown in Fig. 2. The nucleotide sequence surrounding the proposed translation initiation codon for the binding protein agrees with the Kozak consensus characteristic of an active start codon²⁰. Assuming a start site in this position, the GBP transcript carries a stretch of 145 untranslated nucleotides at the 5' end. The termination codon is followed by 46 nucleotides at the 3' terminus with the polyadenylation signal located 15 nucleotides upstream of the start of the poly(A) tail. The deduced amino-acid sequence of GBP is 516 residues long with an estimated M_r of 57,020. The difference between this and the estimated size of the protein purified from brain (68–71K) is probably due to abnormal migration of the brain protein in SDS-PAGE as a result of its isolation in detergent-containing buffers²¹.

A restriction fragment of the insert hybridized to an RNA species of about 1.8 kb (Fig. 3a), which is in agreement with the size of the insert in pGB-A2. The northern blot analyses also indicate that the protein is expressed only in brain, being expressed more in hippocampus, cerebellum and cerebrum than it is in the brainstem region (Fig. 3a).

Hydropathy analysis (Fig. 3b) of the inferred amino-acid sequence²² shows that the protein contains a hydrophobic core region between residues 130 and 330. This core contains at least four segments of relatively high hydrophobicity and long enough

to be potential transmembrane-spanning regions (TM). TM₂ has the same motif as the TM₂ region of other ligand-gated ion channels^{23–25} (Fig. 3c). The similar structural motif of this region may point to the same functional property for the TM₂ segment, namely the formation of the lining of the channel wall²⁶. In the GBP, TM₁ and TM₂ contain two and one positively charged residues, respectively, and TM₃ and TM₄ have one negatively charged residue (Fig. 2). These charged residues in the transmembrane domains may have a role in the unusual voltage sensitivity of the NMDA receptor.

An N-terminal hydrophobic segment has a sequence that is characteristic of a signal peptide²⁷ (Fig. 2). A short C-terminal hydrophobic region ends with a sequence that is very similar but not identical to the canonical structure for isoprenylation (C-A-A-X, where A is an aliphatic and X is any amino acid)²⁸. A protein sequence bank (NBRF) was searched (FASTA program) and the GBP was found to have a unique sequence. There is no substantial homology between this protein and the kainate/AMPA or the metabotropic glutamate receptor. A sequence in the N-terminal region (residues 82 to 117) of GBP has high degree of homology (45.7% identity) to the N-terminal segment (residues 5–38) of pro- α (I) collagen²⁹. Also, two short segments with some homology between GBP and γ -glutamyl transpeptidase³⁰ were detected. The segment with greatest homology is

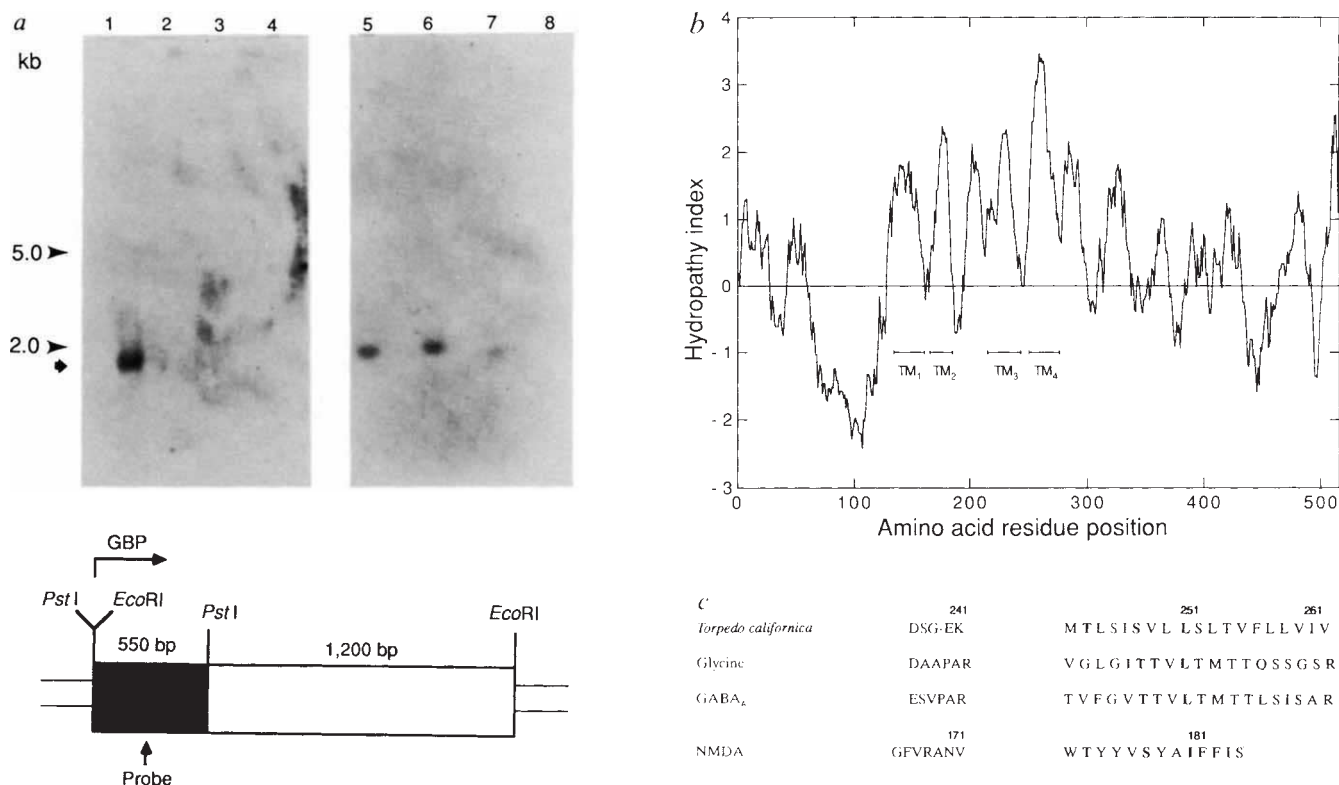


FIG. 3 a, Northern hybridization of a 550-bp PstI restriction fragment of the cDNA for the GBP (bottom) with RNA from 4 different rat tissues (lanes 1, brain; 2, liver; 3, lung; 4, muscle) and 4 different regions of rat brain (lanes 5, hippocampus; 6, cerebellum; 7, cerebrum; 8, brainstem). b, Hydropathy plot of inferred sequence of the GBP (residues 1 to 516) using the algorithm of Kyte and Doolittle²² with a window setting of 15 residues. The proposed transmembrane domains are indicated as TM₁ to TM₄. c, A comparison of the inferred sequences for TM₂ from various subunits of ligand-gated ion channels with the sequence of the putative TM₂ of GBP. The amino-acid sequences for TM₂ shown are those from the α subunit of the nicotinic acetylcholine receptor from *Torpedo californica*²³, the 48K subunit of the glycine receptor²⁵, and the α subunit of the γ -aminobutyric acid (GABA) receptor from rat brain²⁴. The GBP sequence lacks the first negatively charged residue that is characteristic of the other sequences.

METHODS. Each population of RNA was prepared by the guanidine

isothiocyanate-caesium chloride purification method. Total RNA (15 μ g per lane) was size-fractionated on 1% agarose-formaldehyde gels and transferred to Nytran (MSI) membranes. A 550-bp PstI restriction fragment of the GBP cDNA, purified by extracting from a low-melting-temperature agarose gel and labelled with ³²P-dATP by the random primer method (total activity of 5.2×10^8 d.p.m. in 15 ml hybridization buffer) was hybridized to the RNA blot. Prehybridization was at 42 °C for 3 h in 5 $\times$ SSPE (1 $\times$ SSPE is 0.18 M NaCl, 10 mM sodium phosphate, 1 mM EDTA, pH 7.7, 5 $\times$ Denhardt's solution (1 $\times$ Denhardt's is 0.02% Ficoll, 0.02% polyvinylpyrrolidone, 0.02% bovine serum albumin)/0.1% SDS and 50 μ g ml⁻¹ denatured salmon sperm DNA. Hybridization was at 42 °C for 16 h in the same buffer which included 50% formamide. Filters were washed with 1 $\times$ SSPE/0.5% SDS at 37 °C followed by 0.1 $\times$ SSPE/0.1% SDS at 65 °C. Filters were exposed to Kodak X-Omat film with an intensifying screen for 48 h (left panel) or 6 h (right panel).

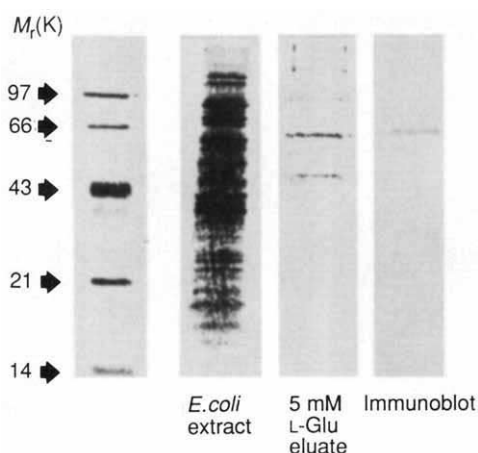


FIG. 4 Expression of the GBP by *E. coli* transfected with plasmid pGB-A2. A protein with $M_r \sim 66$ K was purified from an extract of *E. coli* by affinity chromatography on an L-glutamate-treated ReactiGel column. The western blot of the 5 mM L-glutamate eluate with the antibodies raised against brain GBP revealed staining of the ~ 66 K protein band.

METHODS. The *E. coli* culture carrying pGB-A2 was induced with IPTG. Cells were collected by centrifugation and the pellet was resuspended in 10 mM phosphate buffer, pH 7.4, containing protease inhibitors (see legend to Fig. 1), Tween-20 (0.1% v/v) and Triton X-100 (1% v/v). After sonication (3×10 s), the cell suspension was centrifuged at $46,500g$ for 30 min and the supernatant loaded onto an L-glutamate-treated ReactiGel column. The column was washed with 10 mM potassium phosphate buffer, pH 7.4, containing the protease inhibitors plus 0.1% (v/v) polyoxyethylene-10-tridecyl ether. It was subsequently eluted with 5 mM L-glutamate as described in the legend to Fig. 1. Peak fractions were pooled, dialysed, and electrophoresed on SDS-polyacrylamide gels.

17 amino acids long, is located in TM_2 , and has 60% identity to residues 7 to 24 of γ -glutamyl transpeptidase.

The GBP has two pairs of vicinal cysteine residues in the loop between TM_2 and TM_3 (residues 190, 191) and near the C terminus (residues 431 and 432). The cysteine residues may form disulphide bonds and may play a part in the formation of either agonist or metal binding sites. The inferred protein structure has no canonical sequence for N-linked carbohydrate residues. As at least one form of this protein is glycosylated^{16,17}, we assume that such glycosylation is on either Ser or Thr residues. There are three Ser-Glu or Glu-Ser sequences (residues 101, 102, 433, 434, 506, 507) and six Thr residues followed by charged amino acids. These may be preferred sites for O-glycosylation³¹. Almost all of these sites are in the N- and C-terminal hydrophilic regions; these regions might therefore form two major extracellular domains of the protein.

The cloned protein is believed to correspond to the brain GBP because (1) the induced culture of *E. coli* carrying pGB-A2 yielded a peptide of ~ 60 K that was recognized by the antibodies against the brain GBP and by anti- β -galactosidase antibodies; (2) the same protein was purified from *E. coli* extracts by affinity chromatography of the extracts on a glutamate-treated column (Fig. 4); (3) the eluted protein has an estimated dissociation constant (K_D) for L-glutamate binding ($K_D = 263$ nM) equal to that of the protein purified from rat brain synaptic membranes¹⁶. These results indicate that we have cloned the cDNA for a glutamate-recognition protein from brain which, on the basis of our data, is a subunit of the brain NMDA receptor. □

7. Chapman, A. G. *et al.* *Eur. J. Pharmac.* **149**, 91–96 (1987).
8. Gean, D. W. & Shinnick-Gallagher, A. *Neuropharmacology* **27**, 557–562 (1988).
9. Wellisch, T. *Science* **230**, 681–683 (1987).
10. Choi, D. W. *J. Neurosci.* **7**, 369–379 (1987).
11. Watkins, J. C., Krosgaard-Larsen, P. & Honore, T. *Trends Pharmac. Sci.* **11**, 25–33 (1990).
12. Hollman, M. *et al.* *Nature* **342**, 643–648 (1989).
13. Keinänen, K. *et al.* *Science* **249**, 556–560 (1990).
14. Masu, M. *et al.* *Nature* **349**, 760–765 (1991).
15. Ly, A. M. & Michaelis, E. K. *Biochemistry* **30**, 4307–4316 (1991).
16. Chen, J.-W. *et al.* *J. Biol. Chem.* **263**, 417–426 (1988).
17. Eaton, M. J. *et al.* *J. Biol. Chem.* **265**, 16196–16204 (1990).
18. Mattson, M. P., Wang, H. & Michaelis, E. K. *Brain Res.* (in the press).
19. Short, J. M. *et al.* *Nucleic Acid Res.* **16**, 7583–7600 (1988).
20. Kozak, M. *Nucleic Acids Res.* **12**, 857–872 (1984).
21. Kumar, K. N. *et al.* *J. Biol. Chem.* **266**, 14947–14952 (1991).
22. Kyte, A. M. & Doolittle, F. J. *Molec. Biol.* **157**, 105–132 (1982).
23. Noda, M. *et al.* *Nature* **302**, 528–532 (1983).
24. Schofield, P. R. *et al.* *Nature* **328**, 221–227 (1987).
25. Grenningloh, G. *et al.* *Nature* **328**, 215–220 (1987).
26. Galzi, J. L. *et al.* *A. Rev. Pharmac.* **31**, 37–72 (1991).
27. Heijne, G. V. *Nucleic Acid Res.* **14**, 4683–4690 (1986).
28. Deschenes, R. J. *et al.* *J. Biol. Chem.* **264**, 11865–11873 (1989).
29. Seyer, J. M. & Kang, A. H. *Biochemistry* **16**, 1158–1164 (1977).
30. Tate, S. S. & Meister, A. *Molec. cell. Biochem.* **39**, 357–368 (1981).
31. Fransson, L. A. *Trends Biochem. Sci.* **12**, 406–411 (1987).
32. Sanger, F., Nickleleu, S. & Carlson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).

ACKNOWLEDGEMENTS. We thank J. Boulter of the Salk Institute for the hippocampal cDNA expression library, and M. Michaelis for critically reading the manuscript and L. Kunkle for preparing it. This work was supported by grants from the NIAAA and the ARO, and by the Center for Biomedical Research, University of Kansas.

Dendritic spines as individual neuronal compartments for synaptic Ca^{2+} responses

Wolfgang Müller & John A. Connor

Roche Institute of Molecular Biology, Department of Neurosciences, Roche Research Center, 340 Kingsland Street, Nutley, New Jersey 07110, USA

THE possibility that postsynaptic spines on neuronal dendrites are discrete biochemical compartments for Ca^{2+} -activated processes involved in synaptic plasticity^{1–6} is a widely proposed concept that has eluded experimental demonstration. Using microfluorimetry on CA3 neurons in hippocampal slices, we show here that with weak presynaptic stimulation of associative/commissural fibres, Ca^{2+} accumulates in single postsynaptic spines but not in the parent dendrite. Stronger stimulation also promotes changes in dendrites. The NMDA-receptor antagonist AP-5 blocks changes in Ca^{2+} in spines. Sustained steep Ca^{2+} gradients between single spines and the parent dendrite, often lasting several minutes, develop with repeated stimulation. The observed compartmentalization allows for the specificity^{7,8}, cooperativity⁹ and associativity^{10–14} displayed by memory models such as long-term potentiation.

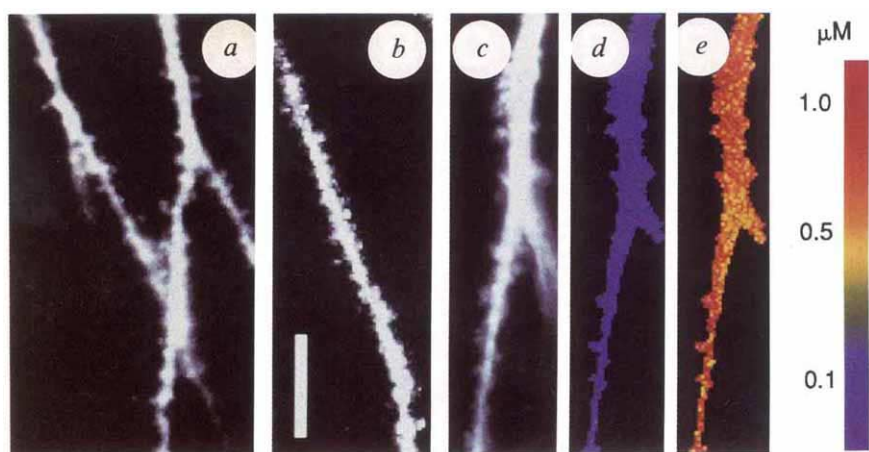
Figure 1a–c demonstrates resolution of the measurement system used, showing the fluorescence of injected Fura-2 in a number of clearly resolvable spines projecting laterally from dendritic shafts of three different neurons. Roughly one out of 20 injected neurons, whose somata were situated within $75 \mu\text{m}$ of the slice surface, gave fluorescent image resolution comparable to that shown in Fig. 1. In these preparations it was straightforward to show large Ca^{2+} increases in both laterally lying spines and the dendrite shaft as a result of stimulating the associative/commissural pathways that form synapses with the distal dendrites of CA3 neurons^{6,15}. Using the dendrite shown in Fig. 1c, panels d and e show Ca^{2+} levels before and during afferent tetanization, respectively (50 shocks of 0.1 ms at 50 Hz). The spines projecting to the left are resolved best and show Ca^{2+} increases from a resting level of 50 nM to values exceeding $1 \mu\text{M}$ during the stimulus train. In some strongly activated spines, Ca^{2+} levels significantly exceeded shaft levels during the stimulus. The Ca^{2+} levels reported during firing (Fig. 1e) are the integrated input of pulsatile Ca^{2+} flux over the acquisition time of the ratio images (≥ 200 ms; methods described in Fig. 1

Received 5 July; accepted 10 September 1991.

1. Fonnum, F. *J. Neurochem.* **42**, 1–11 (1984).
2. Cotman, C. W. *et al.* *Trends Neurosci.* **10**, 273–280 (1980).
3. Balazs, R., Hack, N. & Jorgensen, O. S. *Neurosci. Lett.* **87**, 80–86 (1988).
4. Mattson, M. P., Doug, P. & Kater, S. P. *J. Neurosci.* **8**, 2087–2100 (1988).
5. Monaghan, D. T., Bridges, R. J. & Cotman, C. W. *A. Rev. Pharmac. Toxicol.* **29**, 365–402 (1989).
6. Collingridge, G. L. & Bliss, T. V. P. *Trends Neurosci.* **10**, 288–294 (1987).

FIG. 1 *a, b* and *c*, Fluorescence of microinjected Fura-2 in spine-covered, distal dendritic segments of CA3 pyramidal neurons in living brain slices ($\times 40$ objective; excitation at 380 nm; 200–400 μm apical from the stratum pyramidale; scale bar, 5 μm). *d* and *e*, Ratio images of the dendrite shown in *c* during rest (*d*) and during strong stimulation of associational/commissural fibres in the stratum radiatum (*e*, 50 shocks, 0.1 ms, 50 Hz). Acquisition time for the ratio image pair was 400 ms. Spines projecting to the left show larger increases in Ca^{2+} during the stimulus than most areas of the shaft.

METHODS. Experiments were performed on transverse slices of guinea-pig hippocampus 300 μm thick, placed on a nylon mesh in an interface type slice chamber and continuously perfused with saline (35 $^{\circ}\text{C}$) containing (in mM): NaCl 134, KCl 2, MgSO_4 1.3, KH_2PO_4 1.25, NaHCO_3 27, CaCl_2 2.5, glucose 10; equilibrated with 95% O_2 /5% CO_2 ; pH was 7.4 (refs 17, 19). Drugs were applied to the perfusion saline (DL-2-amino-5-phosphonovaleic acid, AP-5, 50–100 μM and picrotoxin, 5–10 μM ; both from Sigma). Individual CA3 pyramidal neurons were impaled with microelectrodes whose tips contained 12 mM Fura-2 (potassium salt, Molecular Probes) and 200 mM KCl. Electrode barrels were filled with 1–3 M KCl; total resistance was 200–400 M Ω . Neurons with resting potentials less than -60 mV were rejected. Passing iontophoretic current (0.5–1 nA) for 5–20 min achieved 250–500 μM concentrations of Fura-2 in the cells. Cells were imaged from the top surface of the slice using an upright microscope (Zeiss Axioskop) and long distance $\times 20$ and $\times 40$ objectives. In combination with additional optical magnifiers, up to $80\times$ magnification was achieved. A charge-coupled device camera system (Photometrics)^{20,21} was used to acquire digitized images of Fura-2. Calcium concentrations were determined from background-corrected image pairs at 360 nm and 380 nm, or 350 nm and 380 nm excitation using the ratio method^{22,23}. Exposure times were 50–200 ms per



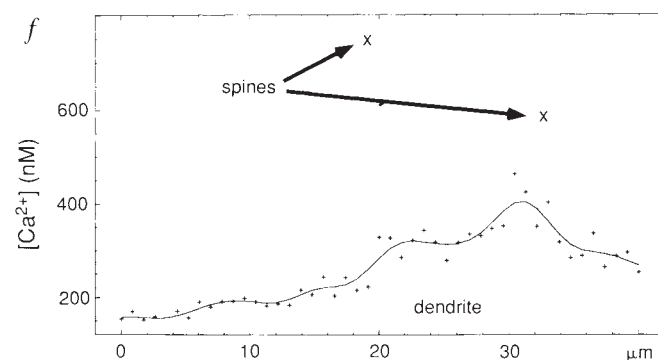
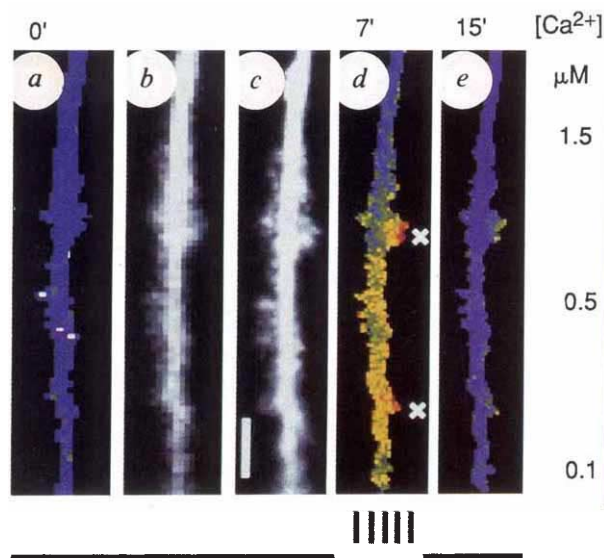
image, giving total acquisition times for a wavelength pair of 200 to 500 ms. The stability of autofluorescence in the slice was monitored routinely during an experiment at several sites around the Fura-2-filled neuron. Data were discarded when low signal-to-background ratio precluded Ca^{2+} measurements within narrow confidence limits ($<10\%$). Ca^{2+} in spines could be determined with sufficient confidence for 30 min to 2 h before confidence limits were violated owing to loss of fluorescence signal. Loss of signal was largely independent of the amount of ultraviolet irradiation and probably reflects diffusion out into distant cell regions. Presynaptic fibres were fired by repetitive stimulation in the stratum radiatum of the CA3 area using a bipolar electrode constructed from twisted pair 25- μm Pt wires. A digitally controlled pulse generator (Master-8, AMPI, Israel) and a step-up transformer were used to drive the bipolar electrode. A TTL pulse from the imaging computer was used to synchronize presynaptic stimulation with image capture. Results are expressed as false colour images, with red coding for highest Ca^{2+} , graphically and as means $\pm$ s.d.

legend). Depending on the Ca^{2+} handling capabilities of the spines, which are unknown, peak Ca^{2+} levels may be considerably higher.

In addition to the differences between the dendrite shaft and laterally projecting spines, Ca^{2+} increase over the dendrite shaft itself was blotchy. We attribute these local differences to spines that are oriented vertically to the plane of view¹⁶ and so give a signal that is added to the shaft signal.

Where many presynaptic inputs were activated, widespread

Ca^{2+} increases resulted which tended to obscure spine/dendrite shaft differences. Still, on average, the differences were considerable, 1.3 ± 0.5 μM for activated spines versus 370 ± 120 nM in the parent dendrites ($n=18$). Where only a few presynaptic fibres were activated, the spine-dendrite differences were much greater because the dendrite shaft was not flooded with Ca^{2+} . A portion of the shaft Ca^{2+} might enter through voltage-gated channels, as we have previously shown using direct depolarization¹⁷.



objective, Ca^{2+} is uniformly low throughout the dendrite. *b* and *c*, Fluorescence images taken with the $\times 20$ objective and the $\times 40$ objective. In both fluorescence images, taken at 380 nm excitation, laterally projecting spines can be recognized, but they are much clearer with the $\times 40$. Scale bar, 10 μm . *d*, Ca^{2+} levels during the seventh train of orthodromic stimulation of associational/commissural fibres in the stratum radiatum (50 shocks, 0.1 ms at 50 Hz). Ca^{2+} rises to high levels in a few spines marked by crosses and to intermediate levels in the lower part of the densely spine-covered dendrite. *e*, Ca^{2+} levels 8 min after the train of *d*. *f*, Graphical profile of the intradendritic Ca^{2+} concentration along the dendrite and in the responding spines (crosses), extracted from the ratio image depicted in *d*.

FIG. 2 Ca^{2+} response in a distal apical dendrite of a CA3 pyramidal neuron to associational/commissural stimulation (50 shocks, 0.1 ms, 50 Hz). *a*, In control conditions, taken during screening of the dendritic tree with a $\times 20$

FIG. 3 AP-5 sensitive and insensitive Ca^{2+} accumulations in response to weak and strong orthodromic stimulation (50 shocks, 0.1 ms at 50 Hz). *a*, Resting Ca^{2+} levels are evenly low throughout the imaged segment of the dendritic tree. *b*, Weak orthodromic stimulation evoked a Ca^{2+} accumulation restricted to one site. *c*, Bath-applied AP-5 (100 μM) almost completely blocked this Ca^{2+} response. *d*, Stronger stimulation evoked a larger Ca^{2+} accumulation at the same site and recruited a second site and, in addition, a widespread intermediate rise of Ca^{2+} . *e*, Bath-applied AP-5 (100 μM) blocked the two local Ca^{2+} accumulations but only marginally reduced the widespread Ca^{2+} response. *f*, Ca^{2+} concentration profiles, through top right and bottom left dendrites and including the strongly responding sites, extracted from the ratio images of *d* and *e*.

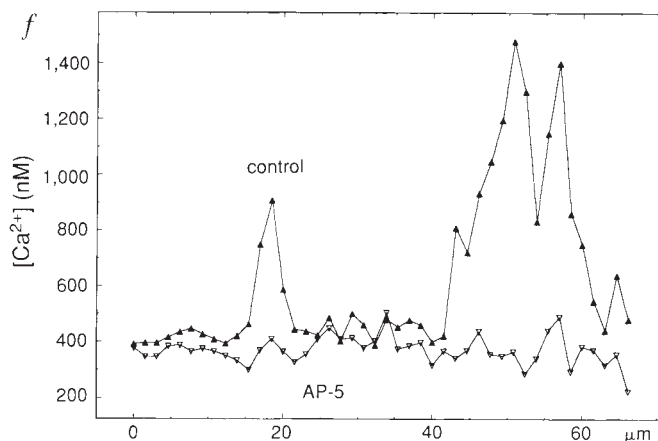
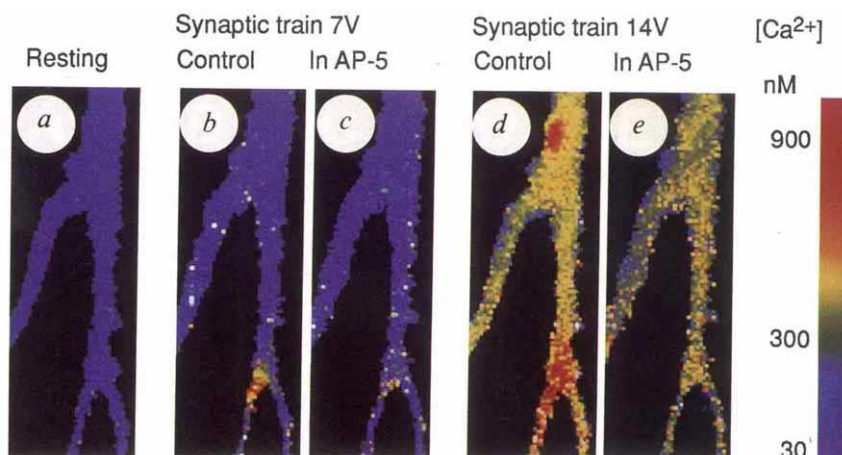
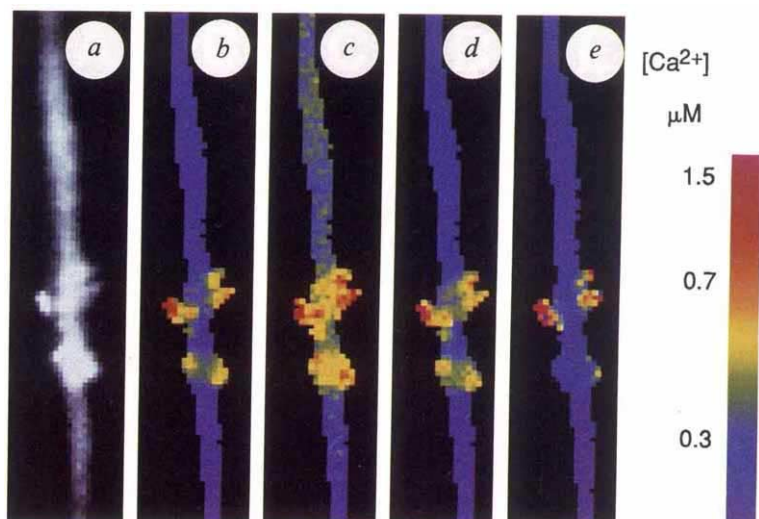


Figure 3 illustrates the blockade of localized increases in Ca^{2+} associated with spines by bath application of the *N*-methyl-D-aspartate (NMDA)-receptor antagonist, DL-2-amino-5-phosphonovaleric acid (AP-5, 50–100 μM ; $n = 5$). Resting Ca^{2+} distribution is shown in Fig. 3*a*. During orthodromic stimulation at low intensity (50 shocks of 0.1 ms, 7 V at 50 Hz), a restricted site of Ca^{2+} accumulation was evident at the lower dendritic fork (Fig. 3*b*). Stronger stimulation (50 shocks of 0.1 ms, 14 V at 50 Hz) evoked a larger Ca^{2+} accumulation at this site and

In light of these results, the main experimental limitations were finding dendrites near the surface of the slice and obtaining selective stimulation of a few spines. Finding the best conditions was facilitated by screening preparations using a $\times 20$ objective and various stimulation intensities before switching to a $\times 40$ objective. Using this procedure, control ratio images of unstimulated dendrites were obtained in many cases using a $\times 20$ objective (Fig. 2*a*). Comparative single-wavelength fluorescence images obtained with the $\times 20$ and the $\times 40$ objectives are shown in Fig. 2*b* and *c*. Resting levels of Ca^{2+} in spines (70 ± 50 nM; $n = 100$) were indistinguishable from dendritic levels in unstimulated cells (76 ± 65 nM; $n = 15$).

Figure 2*d* shows Ca^{2+} accumulations in the dendrite segment during stimulation (50 Hz, 0.1-ms shocks for 1 s). Ca^{2+} levels in two sets of spines (denoted by crosses) increased sharply while smaller, diffuse increases occurred in the dendrite shaft. Apparently the active presynaptic fibres impinged on the central part of the dendrite, as spiny areas above and below showed only small changes during stimulation. Figure 2*f* gives the Ca^{2+} profile along the dendrite extracted from Fig. 2*d*. The spine values of Ca^{2+} (crosses) are far higher than those in the dendritic shaft, even those in the segments underlying the spines. Figure 2*e* illustrates the Ca^{2+} levels at the same location 8 min after the record was made for Fig. 2*d*. Seven intervening stimulus trains had been applied and the most responsive spines (upper cross) still showed slightly increased Ca^{2+} concentrations (see also Fig. 4) while other spines showed recovery to dendrite shaft levels.

FIG. 4 Synaptically induced, sustained Ca^{2+} accumulations in spines on a distal apical dendrite of a CA3 neuron. *a*, Fluorescence image made using 350 nm excitation where fluorescence increases with Ca^{2+} level. Spines stand out brightly because spine Ca^{2+} levels are well above dendritic levels. *b*, Ca^{2+} levels after several well spaced stimulus trains (see text) showing 5 sites of elevated Ca^{2+} levels ~ 1 μM above the dendritic level. *c*, Ca^{2+} levels during the ninth stimulation. Ca^{2+} rises further in the spines and widespread in the dendrite, thus the gradient between the spines and the dendrite is roughly maintained. *d*, Intracellular Ca^{2+} levels, imaged two seconds after *c*, have recovered to the levels observed before the ninth stimulation. *e*, Ca^{2+} levels 4 min after the ninth train. No further stimulation was applied and Ca^{2+} has recovered to almost dendritic levels in the lower set of spines and in the spine projecting onto the dendrite while Ca^{2+} remained high in the upper laterally projecting spines.



recruited a second site (Fig. 3d). In addition, there was a widespread intermediate Ca^{2+} rise in the rest of the dendrite. Bath-applied AP-5 blocked the development of these hot spots when the same stimuli were applied (Fig. 3c and e). Figure 3f shows Ca^{2+} levels plotted along a line down the right fork of the dendrite encompassing the two hot-spine regions. AP-5 totally blocked the localized increases, but left the shaft increases largely unchanged. In five separate neurons the addition of AP-5 completely blocked localized spine increases and reduced the widespread accumulations in the dendrite by 10–50%. Localized Ca^{2+} responses started to recover within 20 min of AP-5 wash ($n = 3$).

With repeated stimulation, the time course of Ca^{2+} recovery in the most highly activated spines slowed, whereas the dendrite and remaining spines recovered within the usual 3–10 s. This resulted in a sustained Ca^{2+} gradient between a previously activated spine and its parent dendrite ($n = 28$ in 5 cells). The GABA_A antagonist picrotoxin (5–10 μM) potentiated Ca^{2+} responses during stimulation and seemed to facilitate the induction of long-lasting spine–dendrite gradients. Figure 4 shows an example of large maintained gradients between spines and the parent dendrite shaft. A fluorescence image (excitation at 350 nm) of the dendrite segment is shown in Fig. 4a. At the time the segment was first examined in the $\times 40$ objective field, eight stimulus trains (50 Hz, 1 s) had been delivered to the stratum radiatum, one while the bath contained picrotoxin (10 μM) and seven while it was being washed out. These trains had been spaced over a period of 50 min. The record in Fig. 4b was made 10 min after the eighth stimulus train and shows high Ca^{2+} levels remaining in two sets of spines. Figure 4c shows Ca^{2+} levels during the ninth stimulus train. There was a further rise of Ca^{2+} in the spines and a widespread intermediate Ca^{2+} accumulation in the parent dendrite. In the four minutes following this stimulation no further stimulation was applied and Ca^{2+} recovered almost to dendritic levels in the lower set of spines (~ 250 nM) while remaining at high levels (900–1,200 nM) in the upper set of spines.

These *in situ* determinations of Ca^{2+} accumulations in spines and their parent dendrites responding to orthodromic stimulation are evidence for the long-standing hypothesis that spines are a discrete, NMDA receptor-coupled compartment for Ca^{2+} -activated processes. The results show that there is a high degree of stimulus specificity with limited presynaptic input; under appropriate conditions, the Ca^{2+} levels in only a few spines respond to orthodromic stimulation, and neighbouring spines and parent dendrite are largely unaffected. Moreover, secondary processes, as yet unidentified, can maintain elevated Ca^{2+} levels in spines for timespans that are orders of magnitude longer than the presynaptic activity, perhaps enabling the spine to function as an information-processing and storage unit. NMDA-induced sustained Ca^{2+} elevations have been observed in the dendritic shafts of acutely isolated hippocampal neurons¹⁸, but the degree of spatial partitioning of the Ca^{2+} -changes is far greater here, probably as a result of the spine geometry limiting Ca^{2+} diffusion from the spine (absent in the isolated cell preparation) to the dendritic shaft. □

16. Harris, K. M. & Stevens, J. K. *J. Neurosci.* **9**, 2982–2997 (1988).
17. Müller, W. & Connor, J. A. *Neuron* **6**, 901–905 (1991).
18. Connor, J. A., Wadman, W. J., Hockberger, P. E. & Wong, R. K. S. *Science* **240**, 649–653 (1988).
19. Müller, W., Misgeld, U. & Heinemann, U. *Expl Brain Res.* **72**, 287–298 (1988).
20. Tank, D. W., Sugimori, M., Connor, J. A. & Llinas, R. *Science* **242**, 773–777 (1988).
21. Regehr, W. G., Connor, J. A. & Tank, D. W. *Nature* **341**, 533–536 (1989).
22. Grynkiewicz, G., Poenie, M. & Tsien, R. Y. *J. biol. Chem.* **260**, 3440–3450 (1985).
23. Tsien, R. Y. *A. Rev. Neurosci.* **12**, 227–253 (1989).

ACKNOWLEDGEMENTS. We thank D. Linden for comments on the manuscript, D. Tank for data acquisition software, and R. Sun and M. Smeyne for technical assistance. W.M. was supported in part by the Deutsche Forschungsgemeinschaft.

Independent regulation of calcium revealed by imaging dendritic spines

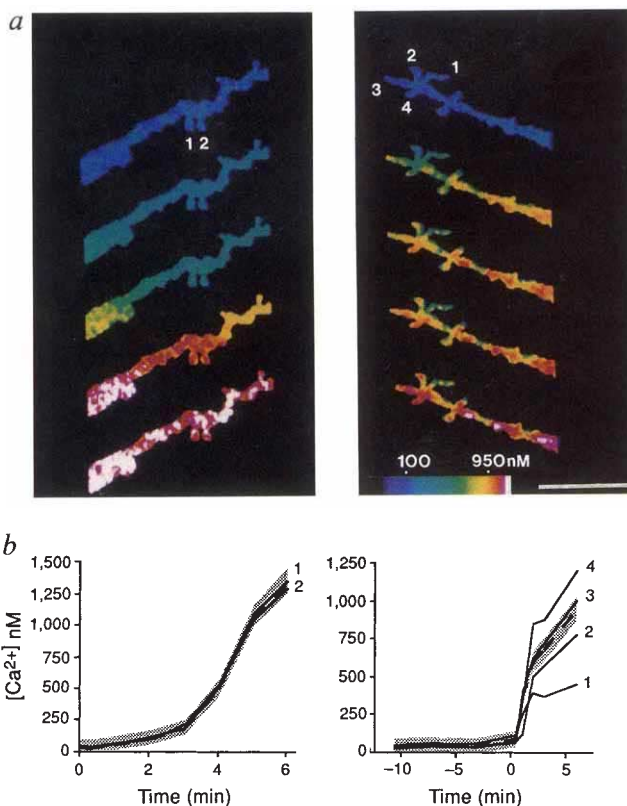
Peter B. Guthrie*, Menahem Segal† & S. B. Kater*‡

* Program in Neuronal Growth and Development, Department of Anatomy and Neurobiology, Colorado State University, Fort Collins, Colorado, USA

† Center for Neuroscience, Weizmann Institute of Science, Rehovot, Israel

‡ To whom correspondence should be addressed

THE dendritic spine is a basic structural unit of neuronal organization. It is assumed to be a primary locus of synaptic plasticity, and to undergo long-term morphological and functional changes^{1–6}, at least some of which are regulated by intracellular calcium concentrations^{7–11}. It is known that physiological stimuli can cause marked increases in intracellular calcium levels in hippocampal dendritic shafts^{12,13}, but it is completely unknown to what extent such changes in the dendrites would also be seen by calcium-sensing structures within spines. Will calcium levels in all spines change in parallel with the dendrite or will there be a heterogeneous response? This study, through direct visualization and measurement of intracellular calcium concentrations in individual living spines, demonstrates that experimentally evoked changes in calcium concentrations in the dendritic shaft ($[\text{Ca}^{2+}]_d$)



Received 10 July; accepted 17 September 1991.

1. Dingledine, R. *J. Physiol.* **343**, 385–405 (1983).
2. Jahr, C. E. & Stevens, C. F. *Nature* **325**, 522–525 (1987).
3. Mayer, M. L. & Westbrook, G. L. *J. Physiol.* **394**, 501–528 (1987).
4. Lynch, G., Larson, J., Kelso, S., Barrionuevo, G. & Schottler, F. *Nature* **305**, 719–721 (1983).
5. Collingridge, G. L., Kehl, S. L. & McLennan, H. *J. Physiol.* **334**, 33–46 (1983).
6. Zalutsky, R. A. & Nicoll, R. A. *Science* **248**, 1619–1624 (1990).
7. Kelso, S. R., Ganong, A. H. & Brown, T. H. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5326–5330 (1986).
8. Malinow, R. *Science* **252**, 722–724 (1991).
9. McNaughton, B. L., Douglas, R. M. & Godard, G. V. *Brain Res.* **157**, 277–293 (1978).
10. Levy, W. B. & Steward, O. *Brain Res.* **175**, 233–245 (1979).
11. Barrionuevo, G. & Brown, T. H. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7347–7351 (1983).
12. Madison, D. V., Malenka, R. C. & Nicoll, R. A. *A. Rev. Neurosci.* **14**, 379–397 (1991).
13. Brown, T. H., Kairiss, E. W. & Kennan, C. L. *A. Rev. Neurosci.* **13**, 475–511 (1990).
14. Cotman, C. W., Monaghan, D. T. & Ganong, A. H. *A. Rev. Neurosci.* **11**, 61–80 (1988).
15. Harris, E. W. & Cotman, C. W. *Neurosci. Lett.* **70**, 132–137 (1986).

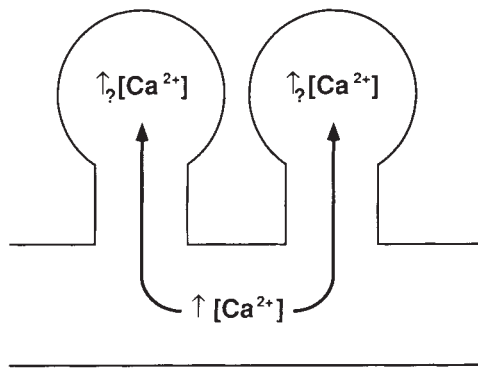


FIG. 1 Schematic of the model being tested. Will calcium concentrations within all spines increase in parallel with increasing dendritic calcium concentrations, or will some spines be protected from increases in dendritic calcium levels?

are frequently not paralleled in the spine ($[Ca^{2+}]_s$). This isolation is not caused by a physical diffusion barrier. This report provides, to our knowledge, the first direct demonstration of autonomous spine function.

Despite the central role of the spine in neuronal communication and the extensive theoretical modelling involving the spine^{5,14-18}, this organelle has actually never been visualized in living neurons. Advances in optical methods recently applied to living neural tissue¹⁹⁻²² have now provided us with the ability to measure properties of dendritic spines directly. Individual pyramidal neurons in the CA1 region of hippocampal slices were injected with the fluorescent calcium indicator Fura-2 (Fig. 2a) and were viewed with a $\times 100$ objective on an inverted microscope (Fig. 2b). Clearly discernible spines were seen in

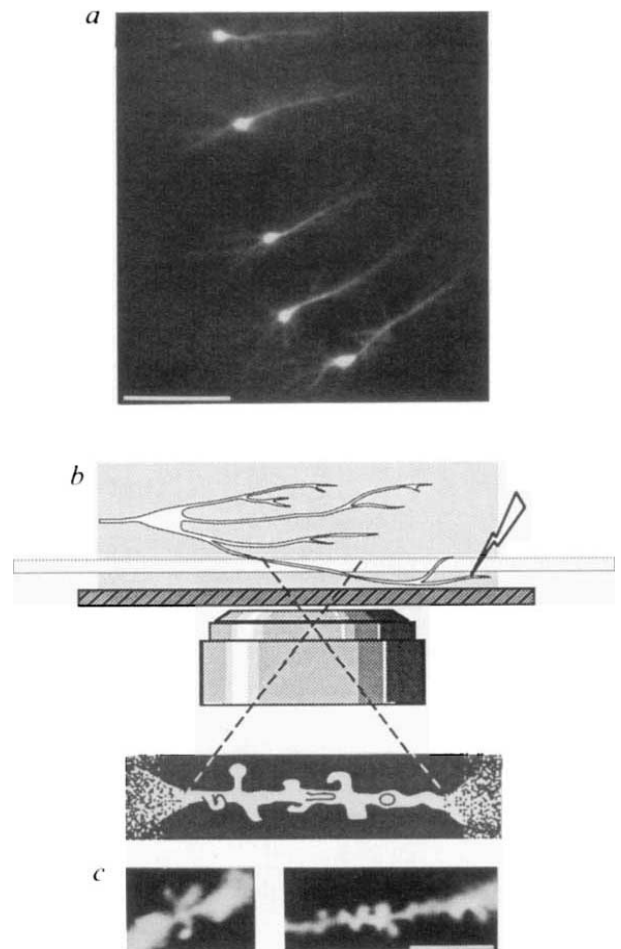


FIG. 2 Fura-2 fluorescence in injected pyramidal neurons in a 250 μm hippocampal slice. *a*, Low-power (Nikon $\times 20$ Fluorite Objective) image of five injected CA1 neurons in a single hippocampal slice. Extensive labelling of dendritic processes is clearly seen. Scale bar, 50 μm . *b*, Schematic of the experimental preparation. Hippocampal slices were prepared and individual CA1 neurons injected with Fura-2. The slices were placed on a No. 0 coverslip with the injected neurons on the bottom of the slice. An oil immersion objective (Zeiss NeoFluor $\times 100$) was used to visualize Fura-2 filled processes within 5–10 μm of the bottom surface of the slice. Processes showing clearly identifiable spines were then used for experimentation. A distant region of the same dendrite (30–80 μm distant) was located. Unfiltered mercury light was focused ($< 3 \mu\text{m}$ diameter) on this distant region of the dendrite for 10–30 s (lightning bolt on diagram). The local, photoinduced damage to the dendritic membrane resulted in a spatially restricted influx of calcium which spread, in a slow wave-like fashion, along the dendrite to the region where spines had been identified. The bottom section of *b* shows a schematic drawing indicating the variety of spine morphologies and orientations observed. The spines radiate in all directions from the dendritic shaft, conferring some additional limitations on visualization. *c*, High-power (Zeiss $\times 100$ NeoFluor Objective) Fura-2 fluorescent (380 nm excitation) image of dendrites and dendritic spines. The image contrast has been computer enhanced for the sake of clarity. Scale bar, 10 μm .

METHODS. Adult rats (150–250 g) were decapitated; their brains were removed and placed in cold carbonate-buffered Krebs solution. The hippocampus was removed and sliced on a tissue slicer at 200–250 μm . Slices were placed in an interface recording chamber where they were exposed to humidified 95% O_2 /5% CO_2 gas mixture, and allowed to recuperate for 1–2 h before recording. Fura-2-filled micropipettes (12 mM; tip resistance 100–200 $\text{M}\Omega$) were used to impale cells in the pyramidal layer of CA1 region. Cells were iontophoretically loaded with Fura-2 for 5–12 min each. Routinely, three to five cells were filled in each slice. Slices were transferred to HEPES-buffered medium at room temperature, pH 7.4, placed on a glass cover slip (size 0) and visualized in an inverted Zeiss IM35 microscope with an intensified CCD camera (Quantex). Slices could be viable in these conditions (judged by recording of cellular activity and population reactivity to afferent stimulation) for at least an hour.

FIG. 3 Measurements of intracellular free calcium concentration in dendrites and spines. *a*, Two series of calcium maps, each of a dendrite and its associated spines before (top image) and following (the four lower images) irradiation of the dendritic shaft (out of the field). Left, times following irradiation for the images are (from top): control; 1 min 10 s; 2 min; 3 min; 6 min. Calcium levels are indicated by colour with cool colours (blue and green) indicating low concentrations, and hotter colours (reds and white) indicating higher concentrations. Irradiation resulted in a wave of increased calcium concentration which progressed down the dendritic shaft and into the spines. $[Ca^{2+}]_s$ in the two spines (1, 2) paralleled $[Ca^{2+}]_d$ over a wide range of values. Right, times following irradiation for the images are (from the top): control; 2 min; 4 min; 5 min; 6 min. The four spines (1, 2, 3, 4) did not all parallel $[Ca^{2+}]_d$. Scale bar, 10 μm . *b*, Calcium concentrations in the dendrites and spines from *a*. The dendrite is indicated by the heavy dashed line. The spines are indicated by thin solid lines, with the numbers to the right corresponding to the spines in *a*. The shaded area indicates the criterion region ± 3 s.d. Left, neither spine deviates from the criterion region. Right, spines 1 and 2 deviated (lagged) from the dendrite for three and four consecutive timepoints, respectively. Spine 3 did not deviate from the criterion region. Spine 4 deviated (led) from the dendrite for three consecutive timepoints.

METHODS. Calcium concentrations were determined from the ratio of fluorescent emission intensities at two different excitation wavelengths²² as described earlier³². The observed 350/380 ratio values were used to determine significant deviation of $[Ca^{2+}]_s$ from $[Ca^{2+}]_d$. We measured the difference in 350/380 ratio ($(350_s/380_s) - (350_d/380_d)$) for 4–9 control time points in 29 spines (nine different preparations). The mean and s.d. (0.01 ± 0.05) were used to estimate the normal variation between spine and dendrite calcium levels. Deviation of spine from dendrite was judged significant if the $(350_s/380_s - 350_d/380_d)$ difference fell outside of ± 3 s.d., in the same direction, for two consecutive time points. Requiring two consecutive time points eliminated the possibility of a single aberrant image producing false deviations. Using this criterion, 1/146 (0.7%) spines were significantly different from the parent dendrite at rest. At least 10 measurements could be taken of the same dendritic region without inducing any noticeable photodamage or bleaching.

the living dendrites, but only near the surface of the slice (Fig. 2c). The intensity of spine fluorescence evoked by either 350 nm or 380 nm excitation was at least, and generally greater than, three times background levels, allowing for reliable measurements both of morphology and intracellular free calcium levels. When viewed as unprocessed fluorescent images, spines could be classified as either 'stubby' ($n = 46$), 'thin' ($n = 35$) or 'mushroom' ($n = 64$)⁵ and ranged in length from 0.5 to 2 μm . Clearly, with the limits of optical microscopy, and the three-dimensional nature of the preparation, some of these classifications may not have been correct. Nonetheless, neither spine shape nor spine size correlated with any of the observations reported here.

Calcium levels at rest, measured in both spines and their parent dendritic shafts, were consistently uniform (79 ± 3 nM; mean $\pm$ s.e.m.; $n = 272$; 42 neurons in 22 animals). For results

reported here, minimum criteria were: (1) signal-to-background ratio had to be at least 3.0 for each excitation wavelength for both the spine and the adjacent dendrite; and (2) rest $[\text{Ca}^{2+}]_s$ and $[\text{Ca}^{2+}]_d$ had to be lower than 200 nM. We found no differences in rest $[\text{Ca}^{2+}]_s$ or rest $[\text{Ca}^{2+}]_s/[\text{Ca}^{2+}]_d$ which correlated with morphological classification. Comparison of individual spines with their immediately adjacent parent dendritic shaft revealed a rest level ratio ($[\text{Ca}^{2+}]_s/[\text{Ca}^{2+}]_d$) of 1.04 ± 0.02 ($n = 145$). This provides strong evidence for equilibration between these two compartments at resting calcium levels.

Isolation between compartments, if it exists, would logically be expected to be most important during periods of neuronal activity. It is already known that physiological stimuli relevant to neuronal activity can cause significant elevations of calcium in the dendrites of hippocampal neurons in slice preparations. We have used a method for raising dendritic calcium to similar levels at a distance site without involving calcium channels²³. We evoked experimentally reliable changes in dendritic shaft calcium levels as follows: a region of the Fura-2-loaded dendrite was selected which was distant (20–50 μm) from the site where spine/dendrite measurements were to be made (Fig. 2b). Then, the unfiltered mercury light was focused (<3 μm diameter) on this distant region of the dendrite for 10–30 s. The local, photoinduced damage to the dendritic membrane resulted in a spatially restricted influx of calcium which spread, in a slow, wave-like fashion (~ 0.1 $\mu\text{m s}^{-1}$) along the dendrite to the region where the experiment was to be conducted. This method evoked a rise in dendritic shaft calcium concentrations at the site of measurements which ranged from 0 nM (for very short duration irradiation) to $>2,000$ nM (for longer duration irradiations). The results presented here are restricted to evoked dendritic calcium rises of at least 200 nM, and no more than 1,500 nM. In 68% (50/74) of the spines examined (27 cells in 17 animals), the change in $[\text{Ca}^{2+}]_s$ mimicked the change in $[\text{Ca}^{2+}]_d$, suggesting coupled regulation of $[\text{Ca}^{2+}]_i$ in the spine and the dendrite (Fig. 3). Not all spines were so rigidly linked to their parent dendritic shaft; in 32% (24/74) of the cases, $[\text{Ca}^{2+}]_s$ deviated significantly

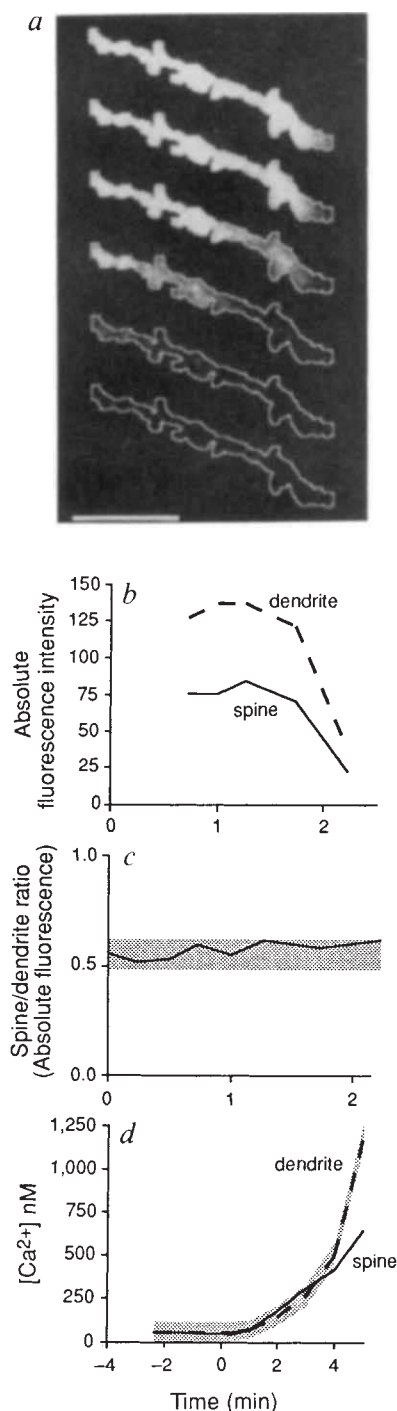


FIG. 4 Diffusion of ions from the dendrite into spines is not limited by a physical barrier. **a**, Fura-2 fluorescent images (30-s intervals) of a dendrite during the arrival of a cobalt quench wave. The boundaries of the dendrite and spines have been outlined. There is no apparent delay in cobalt access to the spines, as indicated by the quenching of Fura-2 fluorescence. Scale bar, 10 μm . **b–d**, Data from an individual spine showing that a spine that deviated in calcium concentrations showed no delay in cobalt access. **b**, Fura-2 fluorescence was measured in the spine and adjacent dendrite region during arrival of the cobalt quench wave. Background images (fully quenched images) were subtracted before making intensity measurements. During a single 30-s interval, the fluorescence intensity in the spine and dendrite were reduced by 69%. During the subsequent time period, fluorescence intensities dropped to background levels. **c**, The ratio of spine fluorescence (F_s) to dendrite fluorescence (F_d) is plotted for the same spine. The shaded region corresponds to the ± 3 s.d. criterion region around the control time period. The ratio F_s/F_d never deviated from the criterion region. **d**, Before the addition of cobalt, the same spine had failed to follow changes in calcium concentrations in the dendrite. The heavy dashed line is $[\text{Ca}^{2+}]_d$; the thin solid line is $[\text{Ca}^{2+}]_s$. The shaded region corresponds to the ± 3 s.d. criterion region. Calcium levels in the spine clearly lagged behind changes in calcium concentrations in the dendrite.

METHODS. Following photoirradiation, CoCl_2 (final concentration 5 mM) was added to the slice. Cobalt quenches Fura-2, leading to disappearance of the fluorescent signal²⁷. Because the cobalt quench analysis required measurement of absolute fluorescence intensity, the ratio of spine fluorescence to dendrite fluorescence (F_s/F_d) was used to control for the different thickness of dendrites and spines. The s.d. of fluorescence intensity ratios, measured in control periods before the cobalt reached the measurement region, was 4%. A significant deviation of F_s from F_d was indicated if the (F_s/F_d) ratio deviated by ≥ 3 s.d. (12%) of the average (F_s/F_d) ratio measured during the control period.

from $[Ca^{2+}]_d$. In 88% (21/24) of these deviating cases, the rise in $[Ca^{2+}]_s$ lagged behind the rise in $[Ca^{2+}]_d$; in 12% (3/24), the rise in $[Ca^{2+}]_s$ preceded that of the $[Ca^{2+}]_d$. Signal-to-background criteria required that the final signal-to-background ratios for all wavelengths for both the spines and the dendrites were no less than three. Using considerably more stringent criteria (minimum signal-to-background of seven), quantitatively similar results were obtained: 5/17 spines (29%) were significantly different from the dendrite.

That these results represent a genuine independence of spines is supported by the fact that virtually identical data on deviation between spines and their parent shafts were obtained using a much less specific method for raising calcium levels. Addition of the potassium channel blocker 4-aminopyridine (4-AP)²⁴⁻²⁶ produced a large, reproducible rise in $[Ca^{2+}]_s$. In 50% (4/8, five preparations), changes in $[Ca^{2+}]_s$ did not parallel changes in $[Ca^{2+}]_d$: of these, 50% (2/4) lagged behind the dendrite, and 50% (2/4) preceded the dendrite. These results support the idea of independent regulation of calcium by dendritic spines and imply that spines can behave as autonomous calcium compartments. In addition, these data suggest that relatively normal spines, with intact presynaptic inputs, were being observed, as 4-AP has been suggested to act primarily at the presynaptic terminal in the hippocampal slice²⁶. It is possible that those spines which preceded the dendrite were being directly stimulated by active presynaptic terminals. Those spines which lagged behind the dendrite might not have been apposed to active presynaptic terminals (or perhaps had no voltage-gated calcium channels), and might represent nonstimulated spines being isolated from changes in calcium concentrations in the adjacent dendrite.

It is of interest that spines were more likely to deviate from the shaft for higher levels of $[Ca^{2+}]_d$. In fact, the local photoirradiation procedure produced two populations of calcium changes: one with a mean $\Delta[Ca^{2+}]_d$ of 350 nM and a second with a mean $\Delta[Ca^{2+}]_d$ of 1,000 nM. Only dendrites from the second population showed differential spine responses. Nonetheless, these calcium increases are well within the range shown to occur in hippocampal dendrites during orthodromic stimulation^{12,13}.

The differential rise of $[Ca^{2+}]_s$ in the spine and parent dendrite suggests that calcium levels in the spine are isolated from changes in the parent dendrite. One possible mechanism would be provided if the spine neck served as a physical barrier for the diffusion of elevated calcium from the dendrite to the spine. If this explanation were true, then $[Ca^{2+}]_s$ should be more likely to lag in the longer 'thin' spines than in the shorter 'stubby' spines; in fact, 21% (4/19) of the 'thin' spines lagged, whereas 31% (5/16) of the 'stubby' spines lagged. Therefore, the spine neck is not likely to act as a significant physical barrier to calcium diffusion. In fact, there was no evidence for differential responses between the different morphological classes of spines.

As a direct experimental test of a possible diffusion barrier, cobalt (CoCl₂: final concentration 5 mM), which acts like calcium but does not activate calcium-dependent processes, was added to the slice following photoirradiation. Cobalt quenches Fura-2, leading to disappearance of the fluorescent signal²⁷. A local decrease in fluorescence intensity indicated that cobalt entered the dendrite exclusively at the site of the irradiation. As the wave of decreased fluorescence (indicating the presence of Co²⁺) reached the site of our measurements, we could ask whether the decrease in Fura-2 fluorescence occurred with a delay in the spine with respect to the dendrite (Fig. 4a). We detected no significant barrier in the rate of transfer of cobalt from the shaft into the spine, that is the fluorescent signal disappeared in the spine at the same time as it did in the parent dendrite in 26 of 28 spines examined. In one case, simultaneous cobalt quench of the Fura-2 signal in the spine and the dendrite was observed in the very same spine in which increases in $[Ca^{2+}]_s$ lagged behind $[Ca^{2+}]_d$ (Fig. 4b-d). Even images taken at 6-s

intervals showed no delay in quenching of spine Fura-2; in contrast, images taken as infrequently as at 15-60 s intervals showed clear delays in increased $[Ca^{2+}]_s$ when compared with $[Ca^{2+}]_d$. These data strongly argue against the likelihood that the delayed increases in spine calcium levels described above could be accounted for by a simple physical barrier to diffusion; rather, we favour an explanation implicating independent calcium clearance mechanisms within the spine shaft or head, such as the spine apparatus. These experiments also suggest that the irradiation-induced calcium increases did not result from influx through voltage-gated calcium channels, as addition of cobalt did not result in an immediate decrease in calcium concentrations in regions distant from the site of irradiation.

These results were obtained near the limits of optical microscopy. For example, only spines within several micrometres of the surface of the slice could be imaged. Dendritic spines are considerably smaller than other structures we have previously imaged, leading to potential problems with dimness, autofluorescence from the tissue, and scatter from adjacent regions of the neuron. Autofluorescence was not likely to be a problem as the injected neuron clearly stood out against the dark background of un-injected cells. In addition, background removal was a standard part of the image processing protocol. Scatter of fluorescent signals from the brighter dendritic shaft was also eliminated as a possible artefact as calcium concentrations in the spines showed no correlation with length of the spine (implying that scatter contribution, which should be a function of distance from the source, was negligible). In addition, as stubby spines had the same probability of showing isolation from the parent dendrite as did the longer thin spines, scatter is unlikely to have played a significant part in determining these results. The fact that rest level measurements of $[Ca^{2+}]_s$ and $[Ca^{2+}]_d$ had the same mean and standard deviation also indicates that spine signals can be as trustworthy as those from the shaft.

These experiments demonstrate, for the first time, the ability to measure $[Ca^{2+}]_s$ in living dendritic spines and to measure reliably changes in $[Ca^{2+}]_s$ in these spines. Previous theoretical investigations have proposed that the dendritic spine can regulate free calcium levels independently from the parent dendrite^{6,28-31}. Our results provide the first experimental proof that spines possess independent mechanisms for regulating $[Ca^{2+}]_s$ such that they can be protected from increases in $[Ca^{2+}]_d$. It will be interesting to use similar techniques to determine whether the dendritic shaft is also often protected from elevated calcium concentrations within spines. The observed isolation of $[Ca^{2+}]_s$ from changes in $[Ca^{2+}]_d$ provides a possible mechanism for the selective modification of individual spines during synaptic activity. □

Received 5 August; accepted 30 September 1991.

1. Ffifková, E. & van Harreveld, A. *J. Neurocytol.* **6**, 211-230 (1977).
2. Ffifková, E. *Cell. molec. Neurobiol.* **5**, 47-63 (1985).
3. Pongracz, F. *Neuroscience* **15**, 933-946 (1985).
4. Winslow, R. L., Miller, R. F. & Ogden, T. E. *Proc. natn. Acad. Sci. U.S.A.* **86**, 387-391 (1989).
5. Harris, K. M. & Stevens, J. K. *J. Neurosci.* **9**, 2982-2997 (1989).
6. Wickens, J. *Prog. Neurobiol.* **31**, 507-528 (1988).
7. Ffifková, E., Markham, J. A. & Delay, R. *J. Brain Res.* **266**, 163-168 (1983).
8. Ffifková, E. & Morales, M. *Ann. N.Y. Acad. Sci.* **568**, 131-137 (1989).
9. Malenka, R. C., Kauer, J. A., Zucker, R. S. & Nicoll, R. A. *Science* **242**, 81-84 (1988).
10. Williams, S. & Johnston, D. *Neuron* **3**, 583-588 (1989).
11. Lynch, G., Larson, J., Kelso, S., Barrionuevo, G. & Schottler, F. *Nature* **305**, 719-721 (1983).
12. Regehr, W. G., Connor, J. A. & Tank, D. W. *Nature* **341**, 533-536 (1989).
13. Regehr, W. G. & Tank, D. W. *Nature* **345**, 807-810 (1990).
14. Shepherd, G. M. et al. *Proc. natn. Acad. Sci. U.S.A.* **82**, 2192-2195 (1985).
15. Coss, R. G. & Perkel, D. H. *Behav. Neural Biol.* **44**, 151-185 (1985).
16. Segev, I. & Rall, W. *J. Neurophysiol.* **60**, 499-523 (1988).
17. Miller, J. P., Rall, W. & Rinzel, J. *Brain Res.* **325**, 325-330 (1985).
18. Kawato, M. & Tsukahara, N. *J. theor. Biol.* **103**, 507-522 (1983).
19. Connor, J. A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6179-6183 (1986).
20. Lev-Ram, V. & Grinvald, A. *Biophys. J.* **52**, 571-576 (1987).
21. Smith, S. J., MacDermott, A. B. & Weight, F. F. *Nature* **304**, 350-352 (1983).
22. Tsien, R. Y., Rink, T. J. & Poenie, M. *Cell Calcium* **6**, 145-157 (1985).
23. Bentley, D., Guthrie, P. B. & Kater, S. B. *J. Neurosci.* **11**, 1300-1308 (1991).
24. Segal, M., Rogawski, M. A. & Barker, J. L. *J. Neurosci.* **4**, 604-609 (1984).
25. Segal, M. *Brain Res.* **414**, 285-293 (1987).
26. Buckle, P. J. & Haas, H. L. *J. Physiol., Lond.* **326**, 109-122 (1982).
27. Mattson, M., Murrain, M. P. & Guthrie, P. B. *Dev. Brain Res.* **52**, 201-209 (1990).

28. Gamble, E. & Koch, C. *Science* **236**, 1311–1315 (1987).
 29. Brown, T. H., Ganong, A. H., Kairiss, E. W., Keenan, C. L. & Kelso, S. R. in *Neural Models of Plasticity: Experimental and Theoretical Approaches* (eds Byrne, J. H. & Berry, W. O.) 266–306 (Academic, San Diego, 1989).
 30. Andrews, S. B., Leapman, R. D., Landis, D. M. & Reese, T. S. *Proc. natn. Acad. Sci. U.S.A.* **85**, 1682–1685 (1988).
 31. Holmes, W. R. & Levy, W. B. *J. Neurophysiol.* **63**, 1148–1168 (1990).
 32. Mattson, M. P., Guthrie, P. B. & Kater, S. B. *J. Neurosci. Res.* **20**, 331–345 (1988).

ACKNOWLEDGEMENTS. We thank D. Giddings for assistance with preparation of the manuscript, and acknowledge support from the Alzheimer's Disease and Related Disorders Association and a grant from the National Institutes of Health.

Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1

F. Ralf Bischoff & Herwig Ponstingl*

Project Molecular Biology of Mitosis, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, W-6900 Heidelberg, Germany

THE product of the gene *RCC1* (regulator of chromosome condensation) in a BHK cell line is involved in the control of mitotic events¹. Homologous genes have been found in *Xenopus*², *Drosophila*³ and yeast^{4,5}. A human genomic DNA fragment and complementary DNA that complement a temperature-sensitive mutation of *RCC1* in BHK21 cells^{6,7} encode a protein of relative molecular mass 45,000 (M_r 45K) which is located in the nucleus and binds to chromatin⁸. We have recently isolated a protein from HeLa cells that strongly binds an anti-RCC1 antibody and has the same molecular mass, DNA-binding properties, and amino-acid sequence as the 205 residues already identified⁹. HeLa cell RCC1 is complexed to a protein of M_r 25K (ref. 9). We have shown¹⁰ that this 25K protein has a sequence homologous to the translated reading frame of *TC4*, a cDNA found by screening a human teratocarcinoma cDNA library with oligonucleotides coding for a *ras* consensus sequence¹¹, and that the protein binds GDP and GTP. We have referred to this protein as the Ran protein (*ras*-related nuclear protein). In addition to the fraction of Ran protein complexed to RCC1, a 25-fold molar excess of the protein over RCC1 was found in the nucleoplasm of HeLa cells. Here we show that RCC1 specifically catalyses the exchange of guanine nucleotides on the Ran protein but not on the protein c-Ha-*ras* p21 (*p21^{ras}*).

p21^{ras} and *ras*-related proteins bind GTP and GDP and display weak GTPase activity (see ref. 12 for review). The active form required for interaction with effectors is bound to GTP, and the activation by exchange of bound GDP to GTP is enhanced by control proteins stimulating the release of GDP^{13–15}. Using purified RCC1 protein and the *ran* protein (Ran) we show that RCC1 acts on its partner in the complex in a way resembling the interaction of nucleotide exchange factors with *p21^{ras}*. The exchange of GDP for bound [³H]GDP and GTP for bound [³H]GTP is catalysed by RCC1 in a concentration-dependent manner (Fig. 1). The exchange of GTP for [³H]GDP and GDP for [³H]GTP produced similar results (data not shown). The GDP exchange rate was approximately 1.6-fold higher than the [³H]GTP release on Ran. In the cell, the exchange of GTP for GDP could be driven by the 10-fold higher concentration of GTP over GDP, as expected for Ras and other GTPases¹⁶. In other experiments in which the concentration of Ran·GDP was varied, we calculated a molar activity of about 160 mol min⁻¹ using an RCC1 concentration of 110 pM, and a K_m value of 300 nM for Ran·GDP at 25 °C. The catalysed exchange is specific for the guanine nucleotides and for Ran, it did not occur with RCC1 or *p21^{ras}* or with other nucleotides (Fig. 2). RCC1·Ran complex was used as a source

of RCC1 because we have found that the purified RCC1 monomer is not stable.

The incorporation of [³H]GDP into Ran·GDP and Ran complexed to RCC1 varies with the concentration of Mg²⁺ (Fig. 3). As described for *p21^{ras}*, Mg²⁺ stabilizes the Ran·GDP complex and inhibits exchange. When a catalytic amount of RCC1·Ran was also present, however, exchange occurred. When the Mg²⁺ concentration was lowered from 5 mM to 0.5 μM, or when EDTA was present, the exchange rate on Ran increased dramatically. In contrast, if Ran is complexed to RCC1, no GDP is bound in the absence of Mg²⁺ and addition of the ion causes nucleotide binding. This is due to the dissociation of the RCC1·Ran complex by nucleotide and Mg²⁺ as previously shown¹⁰. The fact that a plateau value was obtained with 0.5 μM Mg²⁺ and 0.8 μM [³H]GDP indicates that, when the concentration of Mg²⁺·GDP is limiting, an equilibrium is reached between nucleotide-free Ran in the RCC1·Ran complex and nucleotide-bound Ran in solution.

Specific stimulation of the otherwise very low GDP/GTP exchange on the Ran protein by active RCC1 resembles the action of Ras-GRF on *p21^{ras}* in brain¹⁴, and of SDC25 on RAS2 in yeast¹³. RCC1 has no influence on *p21^{ras}* (Fig. 2), thus the respective signals are strictly separated, in keeping with the

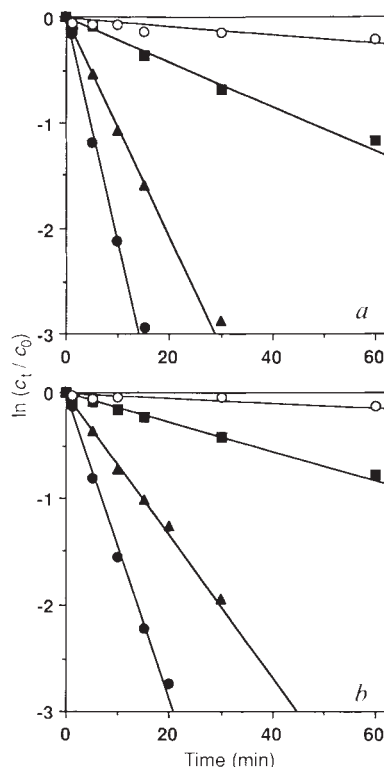


FIG. 1 Catalytic effect of RCC1 on the exchange of Ran-bound [³H]GDP (a) and [³H]GTP (b) by unlabelled guanine nucleotide. The final purification of free nucleoplasmic Ran on hydroxylapatite¹⁰ yielded nucleotide-free Ran in 200 mM phosphate buffer, pH 7.0, containing 1 mM dithiothreitol (DTT) and 1 mM CHAPS; 200 μl of this fraction containing 144 pmol of Ran were incubated with 200 pmol of [³H]GTP (7 Ci mmol⁻¹) or [³H]GDP (10 Ci mmol⁻¹), respectively, for 20 min at 30 °C. The buffer was then changed to 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 5 mM MgCl₂, 1 mM DTT and 0.2 mM CHAPS on a NAP-5 gel filtration column (Pharmacia). The flowthrough was brought to 1.4 ml with Tris-HCl buffer, yielding a concentration of Ran of 86 nM. To 350 μl samples were added 5 μl of unlabelled GDP (a) or GTP (b) to give final concentrations of 140 μM, and 5 μl volumes containing 4.4 (■), 22 (▲) and 44 fmol (●) of RCC1·Ran complex or buffer only (○). After incubation at 30 °C for the periods of time indicated, the ratio of nucleotide bound to Ran at time *t* to the amount bound at time zero (*c_t/c₀*) was determined by vacuum-filtering 50-μl samples through nitrocellulose (Schleicher & Schuell, 0.45 μm), washing with incubation buffer containing 5 mM Mg²⁺, and measuring the protein-bound radioactivity retained by the filter.

* To whom correspondence should be addressed.

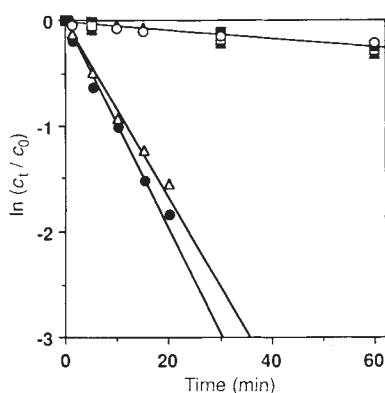


FIG. 2 Specificity of the nucleotide-exchange reaction catalysed by RCC1 for Ran and for guanine nucleotides. Ran·[³H]GDP was prepared as described in Fig. 1. p21^{ras}·[³H]GDP was prepared by incubating 120 pmol of p21^{ras}·GDP¹⁷ in 200 μ l of 20 mM Tris-HCl pH 7.5, 100 mM NaCl, 1 mM DTT and 0.2 mM CHAPS with 200 pmol of [³H]GDP (10 Ci mmol⁻¹) for 20 min at 30 °C. Both samples were supplemented with MgCl₂ and Tris buffer to a final concentration of 5 mM MgCl₂ in 1.4 ml. Samples of 350 μ l containing 30 pmol of Ran·[³H]GDP were incubated with 22 fmol of RCC1-Ran complex (●, ▲), 128 fmol of purified RCC1 (△), or with 5 μ l incubation buffer (○) in the presence of 50 nmol of unlabelled GDP (●, ○, △) or a mixture of 50 nmol each of ATP, CTP, and UTP (▲). Thirty pmol of p21^{ras}·[³H]GDP in 350 μ l of buffer was incubated with 900 fmol of RCC1-Ran complex (■) and with buffer alone (□) in the presence of a 1,000-fold excess of unlabelled GDP. The protein-bound radioactivity was determined from 50 μ l samples as described for Fig. 1.

differing intracellular location of the two proteins. Sequence homology of p21¹², TC4^{11,18} and Ran¹⁰ (Fig. 4) suggests a similar molecular mechanism. Of the 125 residues determined for Ran, including the Ras consensus sequence (positions 65–70 in Fig. 4), only two positions differ from the recently corrected version¹⁸ of the translated cDNA sequence of TC4, which had been found by screening a human teratocarcinoma cDNA library with oligonucleotides encoding the Ras consensus sequence. They are Thr-Thr in Ran (Fig. 4, positions 206–207) as opposed to Asn-Pro in TC4. Thus, Ran and TC4 belong to a group of closely

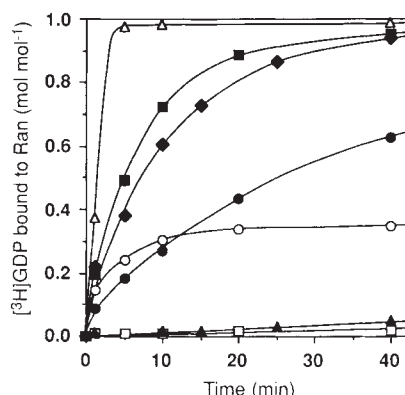


FIG. 3 Effect of Mg²⁺ on nucleotide binding to Ran and to the RCC1-Ran complex. A fraction containing free Ran was incubated with 0.5 mM GDP for 20 min at 30 °C and then gel-filtered on a Sephacryl S-200 column (Pharmacia) in 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1 mM DTT and 0.2 mM CHAPS to remove excess nucleotide and phosphate buffer. Volumes of 250 μ l containing 5.9 pmol of Ran-GDP were incubated with 200 pmol of [³H]GDP (10 Ci mmol⁻¹) in the presence of 1 mM EDTA (■), 0.5 μ M MgCl₂ (●), 5 mM MgCl₂ (▲) and 5 mM MgCl₂ with 4.4 fmol RCC1-Ran (◆). Aliquots of 40 μ l were used to measure protein-bound radioactivity as described in Fig. 1. To measure GDP incorporation using RCC1-Ran complex (○, △, □), 30 pmol of complex were diluted with gel filtration buffer to a final volume of 250 μ l and incubated with 200 pmol of [³H]GDP in the presence of 1 mM EDTA (□), 0.5 μ M MgCl₂ (○) and 5 mM MgCl₂ (△). Samples of 40 μ l were measured, using washing buffers containing Mg²⁺ in concentrations corresponding to the samples.

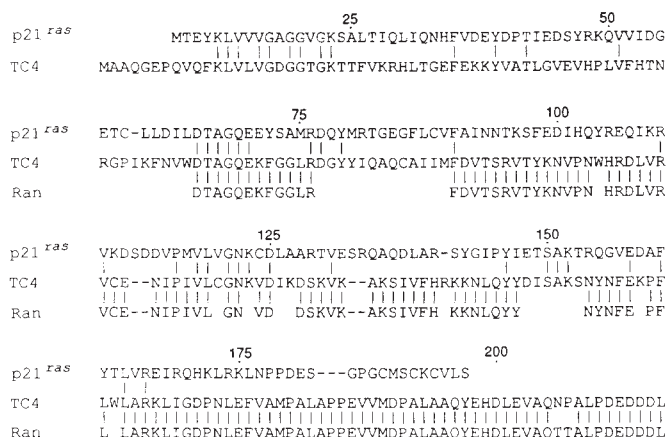


FIG. 4 Sequence comparison of p21^{ras} (ref. 12), TC4¹¹, including a recent correction of a 1-bp frameshift¹⁸, and Ran¹⁰. Vertical lines, identical residues; numbering, corrected TC4. Single-letter amino acid notation.

related sequences that was predicted by Drivas *et al.*¹¹ on the basis of the complex hybridization pattern of TC4 in Southern blots.

By analogy with the *ras* system, we expect the GTP-bound form of Ran to be the active one. The GTP hydrolysis rate of Ran·GTP is very low and we speculate that upon interaction with a thus far unidentified activating protein, Ran becomes a GTPase, as p21^{ras} does upon interaction with GAP (GTPase-activating protein).

In the temperature-sensitive mutant of BHK21 cells, tsBN2, synchronized at the G1/S boundary, shifting to the nonpermissive temperature results in loss of RCC1 protein. As a consequence, p34^{cdc2} kinase is activated and premature chromosome condensation, breakdown of the nuclear envelope and formation of the mitotic spindle occurs independently of the completion of DNA replication^{1,19}. The activation of the p34^{cdc2} kinase after loss of RCC1 requires protein synthesis, indicating that RCC1 could be involved in controlling the expression of a mitotic inducer. In S-phase, the DNA-bound active RCC1 represses premature entry into M phase, possibly by monitoring the replication status of the DNA.

Active RCC1 protein catalyses GDP/GTP exchange on Ran and we propose that this step activates Ran to carry a signal to an effector protein involved in inhibiting the synthesis of the putative mitotic inducer. There would be at least two ways of controlling this signal in the normal cell cycle: RCC1 activity may be regulated by the completion of DNA replication, but there are no data at present indicating variations in activity. Alternatively, RCC1 could be permanently active and Ran·GTP be shut down, for example, by an inhibitor of GDP dissociation or by a specific GTPase activating protein, corresponding to GDI and GAP, respectively, in the *rho*, *rab* and *ras* systems²⁰.

Genes corresponding to mammalian RCC1 and TC4/*ran* have been characterized in *Schizosaccharomyces pombe* by Matsumoto and Beach¹⁸ as being involved in coupling completion of DNA replication to initiation of mitosis. They are designated *pim1* and *spil*. Partial loss of function of the RCC1 homologue *pim1* in mutants was complemented by overexpression of the *ran* homologue *spil*. Assuming identical functions of the respective homologues, this finding can be explained by a reduced nucleotide exchange activity of mutated *pim1* protein, compensated for by an increase in *spil* protein concentration and resulting in a threshold concentration of SP11-GTP sufficient to block premature induction of mitosis. □

Received 7 June; accepted 10 September 1991.

1. Nishitani, H. *et al.* *EMBO J.* **10**, 1555–1564 (1991).

2. Nishitani, H., Kobayashi, H., Ohtsubo, M. & Nishimoto, T. *J. Biochem. Tokyo* **107**, 228–235 (1990).

3. Frasch, M. *EMBO J.* **10**, 1225-1236 (1991).
4. Aebi, M., Clark, M. W., Vijayraghavan, U. & Abelson, J. *Molec. gen. Genet.* **224**, 72-80 (1990).
5. Clark, K. L. & Sprague, Jr. G. F. *Molec. cell. Biol.* **9**, 2682-2694 (1989).
6. Kai, R., Ohtsubo, M., Sekiguchi, M. & Nishimoto, T. *Molec. cell. Biol.* **6**, 2027-2032 (1986).
7. Ohtsubo, M. *et al.* *Genes Dev.* **1**, 585-593 (1987).
8. Ohtsubo, M., Okazaki, H. & Nishimoto, T. *J. Cell. Biol.* **109**, 1389-1397 (1989).
9. Bischoff, F. R., Maier, G., Tiltz, G. & Ponstingl, H. *Proc. natn. Acad. Sci. U.S.A.* **87**, 8617-8621 (1990).
10. Bischoff, F. R. & Ponstingl, H. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
11. Drivas, G. T., Shih, A., Coutavas, E., Rush, M. G. & D'Eustachio, P. *Molec. cell. Biol.* **10**, 1793-1798 (1990).
12. Valencia, A., Chardin, P., Wittinghofer, A. & Sander, C. *Biochemistry* **30**, 4637-4648 (1991).
13. Cr chet, J.-B. *et al.* *Science* **248**, 866-868 (1990).
14. Wolfman, A. & Macara, I. G. *Science* **248**, 67-69 (1990).
15. West, M., Kung, H.-F. & Kamata, T. *FEBS Lett.* **259**, 245-248 (1990).
16. Bourne, H. R., Sanders, D. A. & McCormick, F. *Nature* **349**, 117-127 (1991).
17. John, J., Frech, M. & Wittinghofer, A. *J. biol. Chem.* **263**, 11792-11799 (1988).
18. Matsumoto, T. & Beach, D. *Cell* **66**, 347-360 (1991).
19. Nishimoto, T., Eilen, E. & Basilio, C. *Cell* **15**, 475-483 (1978).
20. Chardin, P. *Cancer Cells* **3**, 117-126 (1991).

ACKNOWLEDGEMENTS. We thank R. H. Himes for discussions, J. Kretschmer for technical assistance, T. Nishimoto for a gift of anti-RCC1-antibody, and A. Wittinghofer for a gift of Ha-ras protein and for discussions. This work was supported by the Deutsche Forschungsgemeinschaft.

A new type of synthetic peptide library for identifying ligand-binding activity

Kit S. Lam, Sydney E. Salmon, Evan M. Hersh, Victor J. Hruby*, Wieslaw M. Kazmierski† & Richard J. Knapp‡

Arizona Cancer Center and Department of Internal Medicine, College of Medicine, Tucson, Arizona 85724, USA

* Department of Chemistry, Faculty of Science, University of Arizona, Tucson, Arizona, USA

† Selectide Corporation, 10900 N. Stallard Place, Tucson, Arizona 85737, USA

OUR aim was to improve techniques for drug development by facilitating the identification of small molecules that bind with high affinity to acceptor molecules (for example, cell-surface receptors, enzymes, antibodies) and so to mimic or block their interaction with the natural ligand^{1,2}. Previously such small molecules have been characterized individually on a serial basis. The systematic synthesis and screening of peptide libraries of defined structure represents a new approach. For relatively small libraries, predetermined sequence variations on solid-phase supports have been used^{3,4}, and large libraries have been produced using a bacteriophage vector into which random oligodeoxynucleotide sequences have been introduced⁵⁻⁸, but these techniques have severe limitations. Here we investigate an alternative approach to synthesis and screening of peptide libraries. Our simple methodology greatly enhances the production and rapid evaluation of random libraries of millions of peptides so that acceptor-binding ligands of high affinity can be rapidly identified and sequenced, on the basis of a 'one-bead, one-peptide' approach.

Our method involves creating a large peptide library consisting of millions of beads, with each bead containing a single peptide and with the complete collection representing the universe of possible random peptides in roughly equimolar proportion. It is clearly not enough to use a random mixture of activated amino acids in a peptide synthesis protocol, because the widely different coupling rates of different amino acids will lead to unequal representation and because each bead will contain a mixture of different peptides. Our solution was to use a 'split synthesis' approach. The first cycle consisted of distributing a pool of resin beads into separate reaction vessels each with a single amino acid, allowing the coupling reactions to go to completion, and then repooling the beads. This cycle was repeated several times to extend the peptide chain (Fig. 1a). In this fashion, each bead should contain only a single peptide species.

We then developed a rapid approach for screening the library to find beads containing peptides able to bind to any particular acceptor molecule. Acceptor molecules were coupled to an enzyme (alkaline phosphatase) or to fluorescein and added in soluble form to the peptide-bead library. Typically, a few beads were intensely stained and were visible to the naked eye and easily seen with a low-power dissecting microscope (bead diameter of 100-200 µm) against a background of colourless, nonreactive beads (Fig. 2). With the aid of tiny forceps coupled to a micromanipulator, the intensely staining beads could be removed for analysis (Fig. 1b). We washed each bead with 8M guanidine hydrochloride to remove the acceptor complex, and then determined the peptide sequence contained on the bead by placing it on a glass filter which was inserted into a peptide microsequencer (model 477A, Applied Biosystems). A library containing several million beads could be screened in 10-15 Petri dishes in an afternoon. Afterwards, it could be washed with 8M guanidine hydrochloride and subsequently reused for screening new acceptors.

Our sequencing studies established that the peptide present on any given single bead is sufficient for unambiguous sequence analysis as each bead sequenced contained 50-200 pmol of peptide (the lower limit of sensitivity of the instrument is in the range of 5 pmol). Furthermore, as measured by preview analysis⁹ for several dozen individual beads from this library, the coupling reactions in the split synthesis procedure were virtually complete as most individual beads displayed peptides that were over 99% pure (range 97-100%). At least three pentapeptide beads were sequenced daily using the microsequencer.

We applied the process to produce a library of millions of pentapeptides and screen it against two well-studied acceptor molecules. Using the split synthesis approach with 19 reaction vessels, we synthesized pentapeptide libraries incorporating all the natural amino acids except for cysteine (for simplicity to eliminate disulphide crosslinking). The random incorporation of 19 amino acids into pentapeptides can produce a total of up to 2,476,099 (19⁵) individual peptides of differing sequence with any one sequence represented on at least one solid-phase resin bead (of course, the number of beads of any given sequence will follow a Poisson distribution, and many multiples of this minimal number of beads are required to assure that the maximum theoretical number of possible peptide entities are approximated in the library). Libraries were readily synthesized in a few days.

We studied a monoclonal antibody against β -endorphin with high affinity ($K_i = 17.5$ nM) for the epitope sequence YGGFL (single-letter amino-acid code). A total of six reactive beads were retrieved from about 2 million beads screened from the pentapeptide library. One peptide ligand sequence retrieved, YGGFQ, had an affinity ($K_i = 15.0$ nM) nearly identical to the native epitope. Two of the other peptide ligands retrieved had K_i s of less than 37 nM (Table 1). These affinities were more than 50-fold better than those obtained for the identical monoclonal antibody by Cwirla *et al.*, who used a phage library method⁷.

TABLE 1 Affinity of anti- β -endorphin ligands: comparison of the natural ligand and peptide ligands identified from a large pentapeptide library

	Sequence	K_i (nM)
Native ligand	YGGFL	17.5 ± 3.2
Peptide ligand	YGGFQ	15.0 ± 1.7
	YGGFA	32.9 ± 2.0
	YGGFT	36.9 ± 7.7
	YGGLS	726 ± 134
	YGALQ	1980 ± 303
	YGGMQ	8780 ± 1500

Affinity constants for the peptide ligands determined with competitive radioligand binding assays using tritiated YGGFL as the standard.

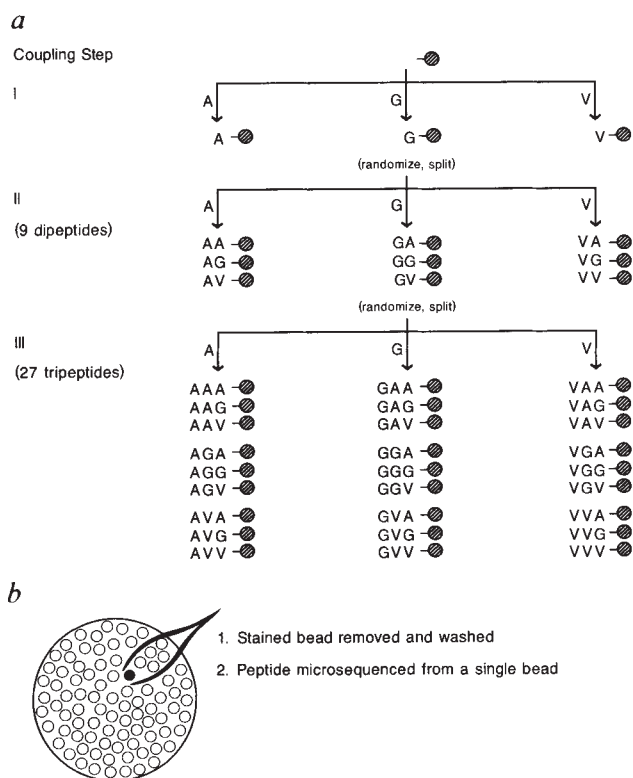


FIG. 1 Steps in peptide library synthesis and screening. *a*, Flow diagram of a simplified example of solid phase 'split synthesis' of tripeptides consisting of alanine (A), glycine (G) and valine (V) using standard solid-phase peptide synthesis methods with Fmoc or Boc chemistry. After each coupling step, the beads from each of the three reaction vessels are combined for randomization and then split again to the three vessels for the next coupling step. After three such steps the 27 possible peptide sequences (3^3) are all represented on separate beads. *b*, A single bead binding an acceptor molecule tagged with an enzyme is identified after being stained by enzymatic reaction on a dye substrate. The stained bead is physically removed from the colourless beads remaining in the library, washed free of the acceptor complex, and subjected to microsequencing. Once identified, the reactive peptide is then synthesized in larger quantities for confirmatory binding and biological studies.

We also used the same pentapeptide bead library to find peptides binding streptavidin (chosen as a receptor-like target). Of 75 reactive beads retrieved, 28 were sequenced and a triplet consensus sequence of HPQ was found in either the amino or carboxy terminus or the central portion of 23 of the recovered pentapeptides (Table 2). The five other peptide ligands sequenced contained the triplet sequence HPM in the pentapeptide. When beads containing LHPQF were synthesized, competitive binding studies established that the HPQ sequence was recognized at the same binding site on streptavidin as the native ligand

TABLE 2 Amino-acid sequences of individual pentapeptide beads that interacted with streptavidin

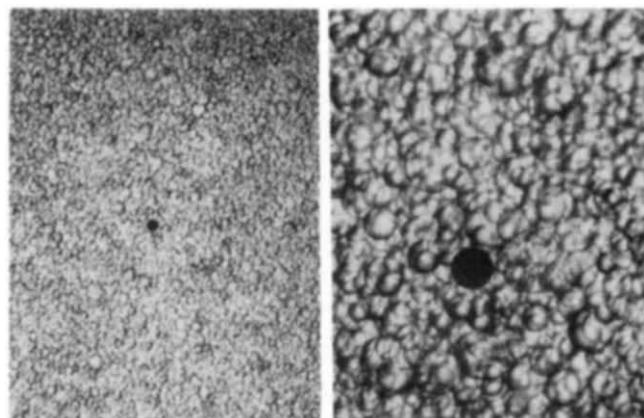
HPQFV	LHPQF	MYHPQ	WNHPM
HPQGP	FHPQG	REHPQ	WTHPM
HQPAG	GHPQN	IQHPQ	VHPMA
	THPQN	GNHPQ	MHPMA (2)
	QHPQG	TVHPQ	
	IHPQG	IGHPQ	
	GHPQG	WMHPQ	
		GAHPQ	
		PLHPQ	
		AIHPQ	
		AAHPQ	
		TPHPQ (2)	

All sequences listed above were found on single beads except for TPHPQ and MHPMA for which two beads were obtained. The first three columns list the sequences found with HPQ located in the amino terminus, central region or carboxy terminus respectively. The fourth column lists the HPM sequences retrieved.

biotin. These studies establish proof of concept for the process, and show that our synthetic libraries can be effectively recycled.

Other investigators have attempted to develop peptide libraries for similar purposes. For example, Geysen and colleagues³ synthesized peptides of known amino-acid sequence on plastic pegs in 96-well plates. This approach permitted the synthesis of several thousand peptides⁴. A related technique using complex instrumentation, photochemistry, and computerized inventory control reported by Fodor *et al.*⁴ permitted synthesis of known arrays of at least 1,024 peptides on an individual microscope slide. Finally, Smith and colleagues pioneered the concept of using a recombinant bacteriophage incorporating random nucleic acids to produce phage displaying millions of random peptides⁵. This approach is innovative but faced with the inherent limitations of synthetic and selection biases of biological systems. Various investigators have used these techniques to identify ligands for several monoclonal antibodies and streptavidin³⁻⁸. On the basis of comparisons with this published information, it is clear that our process is far simpler and more rapid than any of the alternative methods and unlike these methods has identified peptide ligands with affinities virtually identical to those of the native ligands. Additionally, our approach has far greater potential for applying the richness of well-established peptide chemistry to synthesize libraries incorporating D-amino acids or unnatural amino acids as well as specific secondary structures including cyclic peptides. All of this can be accomplished without need to keep records of the synthetic products as our interest is focused just on those peptides which provide a strong interaction signal with the acceptor. On the basis of the combinatorial possibilities and chemical procedures available, it is feasible for us to develop far larger

FIG. 2 Low- and high-power photomicrographs of a peptide ligand library screening in which a reactive (dark) bead stained with the alkaline phosphatase reaction can be easily identified in a background of many thousands of nonreactive (colourless) beads.



libraries of longer or more diverse peptides should they be required for any given application.

We have expanded the applications of our peptide library approach by modifying the synthesis procedure to incorporate cleavable linkers on each bead. After exposure to the cleaving agent, such beads can then release a portion of their peptides into solution for biological assay while still retaining sufficient peptides on the beads for subsequent structure determination.

The one-bead, one-peptide concept and its applications discussed above demonstrate that this approach provides important new tools with which to search for specific ligands of potential diagnostic or therapeutic value. Such information should also enhance fundamental understanding of interactions between ligands and acceptor molecules. □

Received 30 May; accepted 19 September 1991.

1. Hruby, V. J., Al-Obeidi, F. & Kazmiński, W. *Biochem. J.* **268**, 246–262 (1990).
2. Hruby, V. J. & Sharma, S. D. *Curr. Opin. Biotech.* **2**, 599–605 (1991).
3. Geysen, H. M., Melven, R. H. & Barteling, S. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3998–4002 (1984).
4. Fodor, S. P. et al. *Science* **251**, 767–773 (1991).
5. Parmley, S. F. & Smith, G. P. *Gene* **73**, 305–318 (1988).
6. Scott, J. K. & Smith, G. P. *Science* **249**, 386–390 (1990).
7. Cwiria, S. E., Peters, E. A., Barrett, R. W. & Dower, W. J. *Proc. natn. Acad. Sci. U.S.A.* **87**, 6378–6382 (1990).
8. Devlin, J. J., Panganiban, L. C. & Devlin, P. E. *Science* **249**, 404–406 (1990).
9. Niall, H. D., Tregear, G. W. & Jacobs, J. In *Chemistry and Biology of Peptides*, (ed. Meienhofer, J.), 695 (Ann Arbor, Michigan, 1972).

ACKNOWLEDGEMENTS. We thank F. Al-Obeidi, M. Ross and R. Hirschmann for valuable suggestions and E. Lander for comment on the manuscript. K.S.L. is a Special Fellow of the Leukemia Society of America. The peptide synthesis and screening process is referred to as the Selectide Process and is the subject of patent. This work was supported by the National Institutes of Health, the Arizona Disease Control Research Commission and the Selectide Corporation.

Generation and use of synthetic peptide combinatorial libraries for basic research and drug discovery

Richard A. Houghten, Clemencia Pinilla,
Sylvie E. Blondelle, Jon R. Appel, Colette T. Dooley
& Julio H. Cuervo

Torrey Pines Institute for Molecular Studies, 3550 General Atomics Court,
San Diego, California 92121, USA

EXISTING methods for the synthesis and screening of large numbers of peptides are limited by their inability to generate and screen the requisite number (millions) of individual peptides^{1–4} and/or their inability to generate unmodified free peptides in quantities able to interact in solution^{4–8}. We have circumvented these limitations by developing synthetic peptide combinatorial libraries composed of mixtures of free peptides in quantities which can be used directly in virtually all existing assay systems. The screening of these heterogeneous libraries, along with an iterative selection and synthesis process, permits the systematic identification of optimal peptide ligands. Starting with a library composed of more than 34 million hexa-peptides, we present here the precise identification of an antigenic determinant recognized by a monoclonal antibody as well as the straightforward development of new potent antimicrobial peptides.

The initial synthetic peptide combinatorial library (SPCL) prepared and used in this work consisted of six-residue peptide sequences with acetylated N terminals and amidated C terminals. The first two positions in each peptide were individually and specifically defined, whereas the last four positions consisted of equimolar mixtures of 18 of the 20 natural L-amino acids (for ease of synthesis, cysteine and tryptophan were omitted in

this initial library). Such libraries can be generally represented by the sequence Ac-O₂O₂XXXX-NH₂ (where Ac represents acetyl) (see legend to Fig. 1).

Using a competitive enzyme-linked immunosorbent assay (ELISA), each of the 324 different peptide mixtures of the SPCL (Ac-O₂O₂XXXX-NH₂) was assayed to determine its ability to inhibit the interaction of a monoclonal antibody with a larger 13-residue peptide (Ac-YPYDVPDYASLRN-NH₂; single-letter amino-acid code). Of the 324 peptide mixtures examined (Fig. 1), Ac-DVXXXX-NH₂ caused the greatest inhibition of antibody binding (Table 1). Twenty new peptide mixtures were then synthesized in which the third position of the peptide mixture Ac-DVXXXX-NH₂ was defined (Ac-DVOXXXX-NH₂, tryptophan now included in the X positions). Each new peptide mixture contained 6,859 (19³) individual peptides (137,180 in total). The most effective inhibiting peptide mixture was Ac-DVPXXXX-NH₂ (50% inhibitory concentration, IC₅₀ = 41 μM; Table 1b). The above iterative process, which reduces the number of peptide sequences by 20-fold each time it is repeated, was then carried out for the remaining three positions (Table 1, c–e). It should be noted that on defining the fifth position (Ac-DVPDOX-NH₂, Table 1d), the IC₅₀ found for Ac-DVPDYX-NH₂ (0.38 μM) was at least 3,500-fold lower than any of the other 19 peptide mixtures. Also, the peptide mixtures Ac-DVPDXX-NH₂ and Ac-DVPXXX-NH₂ had IC₅₀ values lower than all of the peptide mixtures with the fifth position defined, with the exception of Ac-DVPDYX-NH₂. This clearly

TABLE 1 Identification of the antigenic determinant recognized by monoclonal antibody 19B10

Peptide mixture	IC ₅₀ (μM)	Peptide	IC ₅₀ (μM)
(a)		(e)	
Ac-DVXXXX-NH ₂	250	Ac-DVPDYA-NH ₂	0.03
Ac-DIXXXX-NH ₂	318	Ac-DVPDYS-NH ₂	0.27
Ac-DMXXXX-NH ₂	752	Ac-DVPDYX-NH₂	0.38
Ac-DLXXXX-NH ₂	>1,400	Ac-DVPDYC-NH ₂	0.90
		Ac-DVPDYV-NH ₂	1.10
(b)		Ac-DVPDYT-NH ₂	1.50
Ac-DVPXXX-NH ₂	41	Ac-DVPDYG-NH ₂	1.60
Ac-DVEXXX-NH ₂	146	Ac-DVPDYE-NH ₂	4.06
Ac-DVQXXX-NH ₂	215	Ac-DVPDYI-NH ₂	5.29
Ac-DVXXXX-NH₂	250	Ac-DVPDYM-NH ₂	7.70
Ac-DVMXXX-NH ₂	451	Ac-DVPDYQ-NH ₂	8.18
Ac-DVRXXX-NH ₂	906	Ac-DVPDYH-NH ₂	8.99
Ac-DVAXXX-NH ₂	1,107	Ac-DVPDYL-NH ₂	9.98
Ac-DVCXXX-NH ₂	>1,400	Ac-DVPDYR-NH ₂	10.90
		Ac-DVPDYF-NH ₂	12.02
(c)		Ac-DVPDYN-NH ₂	15.56
Ac-DVPDXX-NH ₂	4.4	Ac-DVPDYK-NH ₂	17.60
Ac-DVPXXX-NH₂	41	Ac-DVPDYY-NH ₂	22.48
Ac-DVPAXX-NH ₂	>1,400	Ac-DVPDYP-NH ₂	26.14
		Ac-DVPDYW-NH ₂	32.14
(d)		Ac-DVPDYD-NH ₂	48.00
Ac-DVPDYX-NH ₂	0.38		
Ac-DVPDXX-NH₂	4.4		
Ac-DVPDAX-NH ₂	>1,400		

The IC₅₀s of the most effective inhibitory peptide mixtures obtained at each iterative step are illustrated for: a, peptide mixtures from the initial screening of the SPCL; b, the third position defined (Ac-DVOXXXX-NH₂); c, the fourth position defined (Ac-DVPOXX-NH₂); d, the fifth position defined (Ac-DVPDOX-NH₂); and e, the sixth position defined (Ac-DVPDYO-NH₂). The IC₅₀ of the peptide mixture derived from the previous iterative step is in bold for comparison. Peptide mixtures were assayed by competitive ELISA (see Fig. 1). The concentration of each peptide mixture necessary to inhibit 50% of the antibody binding to the control peptide on the plate was obtained by serial dilutions of the peptide mixture. The IC₅₀s were calculated using the software GRAPHPAD (ISI, San Diego). The four-step iterative screening and synthesis process takes approximately 4 weeks. This time frame will vary depending on the assay being used and the number of cases moved forward at each iterative step.

illustrates the importance of the presence of specific peptide sequences in each peptide mixture. Among the 20 peptides in which the sixth position was defined, Ac-DVPDYA-NH₂ had the lowest IC₅₀ (0.03 µM; Table 1e). This sequence exactly matches the antigenic determinant found in earlier studies to be recognized by this monoclonal antibody^{3,9,10}. With other screening library procedures⁶⁻⁸, such precise sequence determination, or the identification of different sequences with affinities equal to or exceeding existing sequences, was not accomplished. The results presented here confirm our earlier work in which individual peptides⁹⁻¹¹ or chemically synthesized heterogeneous peptide mixtures^{12,13} were used to establish that each position in a linear antigenic determinant has a specific, quantifiable rank order of importance. The use of this SPCL permits the ready determination of the specific peptide sequence that bound to this antibody out of a total of 34,012,224 possible hexapeptides. Note that no information about the sequence of the antigen or antibody is required to carry out determinations of this kind.

The development of new, potentially useful therapeutic peptides¹⁴⁻¹⁹ requires the synthesis and screening of hundreds to thousands of analogues of an original, often serendipitously discovered active sequence. The potential of SPCLs for the development of new antimicrobial peptides against *Staphylococcus aureus* (Gram-positive bacteria), *Escherichia coli* and

Pseudomonas aeruginosa (Gram-negative bacteria), and the yeast *Candida albicans* was examined in microdilution assays using the same set of 324 peptide mixtures making up the Ac-O₁O₂XXXX-NH₂ peptide library used above. Although many antimicrobial lead peptide sequences derived from this SPCL have been followed in detail (manuscript in preparation), a single example (Ac-RRXXXX-NH₂) found effective against the above microorganisms is presented here. Positions three to six of Ac-RRXXXX-NH₂ were defined using the iterative process described above (the data for *S. aureus* are shown in Table 2). The minimum inhibitory concentrations (MIC) of the 20 individual peptides obtained on defining the sixth position of Ac-RRWXX-NH₂, as well as the C-terminal amide form of the naturally occurring antimicrobial peptide magainin, are shown in Table 3. The hexa-peptide Ac-RRWWCR-NH₂ was the most active of this set. Its MIC against *S. aureus* was 3.2–6.5 µg ml⁻¹. Preliminary data indicate that this peptide is bacteriocidal. The haemolytic activity of Ac-RRWWCR-NH₂ was less than 0.2% at 500 µg ml⁻¹. It is noteworthy that the antistaphylococcal activities of 17 of the 20 sequences were greater than magainin¹⁵.

The use of SPCLs has been illustrated here for both the precise identification of a linear antigenic determinant recognized by a monoclonal antibody and for the development of new, highly effective antimicrobial peptides. A number of other libraries,

FIG. 1 Initial screening of the SPCL (Ac-O₁O₂XXXX-NH₂) for ability to inhibit the binding of monoclonal antibody 19B10. Each of the 324 peptide mixtures of the SPCL was assayed by competitive ELISA for its ability to inhibit the binding of monoclonal antibody 19B10 (ref. 20) to the plate-adsorbed peptide Ac-YPYDVPDYASLRN-NH₂. The individual bar graphs are segregated by first amino acid (O₁), with the individual bars in each graph representing the 18 individual amino acids making up the second position (O₂). The y-axis represents optical density (OD) at 492 nm. The horizontal line in each bar graph represents the average OD of the 324 peptide mixtures. O₁ and O₂ are specific individual amino acids; that is O₁O₂=AA, AD, AE, and so on through to YV, YY, for a total of 324 combinations at positions O₁ and O₂ (18²). Each X position represents an equimolar mixture of the 18 amino acids A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, Y for a total of 104,976 combinations (18⁴). Each of the 324 different peptide mixtures consists of 104,976 individual hexamers, which represent 34,012,224 peptides in total (324 × 104,976). Assuming an average relative molecular mass for Ac-O₁O₂XXXX-NH₂ of 785, then a mixture of 104,976 peptides (18⁴) at a total final concentration of 1.0 mg ml⁻¹ yields a concentration of every peptide in each mixture of 9.53 ng ml⁻¹ (12.1 nmol l⁻¹).

METHODS. The synthetic peptide library was prepared using methylbenzhydrylamine (MBHA) polystyrene resin and standard *t*-Boc chemistry in combination with simultaneous multiple peptide synthesis (SMPS³). A divide, couple and recombine (DCR) process was used to synthesize the XXXX-peptide resin. This process assures equimolarity of the peptides on the resin. Briefly, 18 porous polypropylene packets, each containing 4.65 mmol (5.00 g) of MBHA resin, were coupled to each of the protected *N*-α-*t*-Boc amino acids of interest. All coupling reactions proceeded to completion (>99.5%), as assessed by Gisin's picric acid²¹ or Kaiser's ninhydrin tests²². The resulting resins from each packet were then combined and thoroughly mixed. This resin mixture was separated into 18 portions of equal weight which were placed into porous polypropylene packets, followed by *N*-α-*t*-Boc protecting group removal and neutralization of the resulting amine TFA salts. The resin packets were then reacted with solutions of the individual activated amino acids to yield the 324 dipeptide combinations (18²). The above DCR process was repeated twice more, yielding a final mixture of 104,976 protected tetra-peptide resins (18⁴). This XXXX-resin was divided into 324 aliquots (150 mg each) and placed in numbered, porous polypropylene packets. Synthesis of the next two defined positions was carried out by SMPS. The peptide mixtures were deprotected and cleaved from their respective resins using low-high hydrogen fluoride (HF)²³ as described for individual peptides earlier^{17,24} in a multiple HF cleavage apparatus (Multiple Peptide Systems, San Diego, California). Extraction of the individual peptide mixtures was carried out with H₂O. The time frames for synthesis, and the amounts of each peptide mixture obtained, are the same as described earlier for this number of individual peptides³. The competitive ELISA used is a modification of the direct ELISA technique described previously¹⁰. It differs only in the antibody addition step in which 25 µl each peptide mixture of the SPCL was added, with a fixed dilution of antibody 19B10 (25 µl per well).

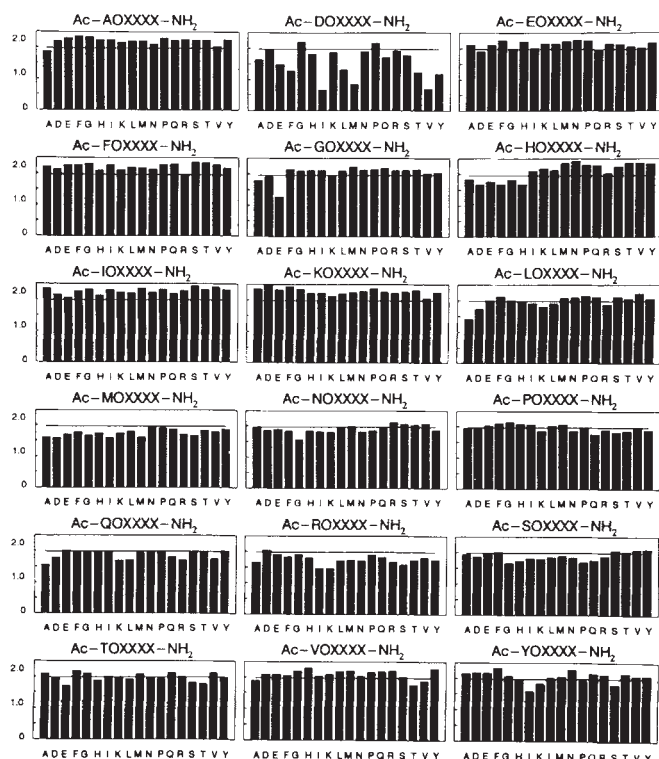


TABLE 2 Antimicrobial activity against *S. aureus* obtained in the iterative process

Peptide mixture	IC ₅₀ ($\mu\text{g ml}^{-1}$)	Peptide mixture	IC ₅₀ ($\mu\text{g ml}^{-1}$)
(a)		(c)	
Ac-RRWXXX-NH ₂	216	Ac-RRWWCX-NH ₂	8.7
Ac-RRYXXX-NH ₂	239	Ac-RRWWX-NH ₂	9.9
Ac-RRRXXX-NH ₂	275	Ac-RRWWRX-NH ₂	9.9
Ac-RRHXXX-NH ₂	286	Ac-RRWWFX-NH ₂	12
Ac-RRCXXX-NH ₂	338	Ac-RRWWIX-NH ₂	14
Ac-RRXXX-NH₂	450	Ac-RRWWXX-NH₂	32
(b)		(d)	
Ac-RRWXXX-NH ₂	36	Ac-RRWWCR-NH ₂	3.4
Ac-RRWXXX-NH ₂	58	Ac-RRWWCW-NH ₂	4.1
Ac-RRWRXX-NH ₂	77	Ac-RRWWCV-NH ₂	4.9
Ac-RRWCXX-NH ₂	88	Ac-RRWWCY-NH ₂	5.4
Ac-RRWLXX-NH ₂	178	Ac-RRWWCK-NH ₂	5.5
Ac-RRWXXX-NH₂	273	Ac-RRWWCX-NH₂	8.7

The five lowest IC₅₀s obtained are illustrated for the peptide mixtures on defining: *a*, the third position (Ac-RRXXX-NH₂); *b*, the fourth position (Ac-RRWXXX-NH₂); *c*, the fifth position (Ac-RRWWCX-NH₂); and *d*, the sixth position (Ac-RRWWCO-NH₂). The IC₅₀ of the peptide mixture derived from the previous iterative step is in bold for comparison. The antimicrobial activity of each peptide mixture against *S. aureus* ATCC 29213 was determined as described earlier¹⁹. Briefly, in 96-well tissue culture plates, peptide mixtures were added to the bacterial suspension ($1-5 \times 10^5$ colony-forming units ml⁻¹) at concentrations derived from serial twofold dilutions ranging from 1.5 mg ml⁻¹ to 2.9 $\mu\text{g ml}^{-1}$. The plates were incubated overnight at 37 °C, and the growth determined at each concentration by the optical density at 620 nm. The relative per cent of growth found for each set of peptide mixtures was consistent in three separate determinations. The IC₅₀s were then calculated using the software program GRAPHPAD (ISI).

TABLE 3 Antimicrobial activity of Ac-RRWWCO-NH₂ against *S. aureus*

Sequence	MIC ($\mu\text{g ml}^{-1}$)	Sequence	MIC ($\mu\text{g ml}^{-1}$)
Ac-RRWWCR-NH ₂	3.2-6.5	Ac-RRWWCA-NH ₂	9-18
Ac-RRWWCV-NH ₂	3.8-7.7	Ac-RRWWCP-NH ₂	10-19
Ac-RRWWCW-NH ₂	4.5-9.0	Ac-RRWWCM-NH ₂	14-27
Ac-RRWWCY-NH ₂	4.7-9.5	Ac-RRWWCL-NH ₂	14-27
Ac-RRWWCK-NH ₂	4.8-9.6	Ac-RRWWCI-NH ₂	17-34
Ac-RRWWCT-NH ₂	4.9-10	Ac-RRWWCF-NH ₂	17-34
Ac-RRWWCH-NH ₂	5.5-11	Ac-RRWWCC-NH ₂	19-38
Ac-RRWWCQ-NH ₂	6-12	Magainin-II-NH₂	32-64
Ac-RRWWCS-NH ₂	7-14	Ac-RRWWCE-NH ₂	>250
Ac-RRWWCX-NH₂	7-14	Ac-RRWWCG-NH ₂	>500
Ac-RRWWCN-NH ₂	8-16	Ac-RRWWCD-NH ₂	>1,000

The MICs for the 20 peptides in which the sixth position is defined (Ac-RRWWCO-NH₂) are shown. The MIC is defined as the lowest concentration of peptide at which no growth is detected after 21 h incubation at 37 °C. The MICs of the previous peptide mixture and magainin-II are bold for comparison.

such as one composed entirely of D-amino acids, have been prepared which in total permit the systematic screening of hundreds of millions of peptides. A fundamental feature of SPCLs is that free peptides can be generated and used in solution in virtually all existing assay systems at a concentration of each peptide most applicable to the assay. This approach has also been successfully used in radio-receptor assays (opioid peptides) and plaque inhibition assays (human immunodeficiency virus (HIV-1) and herpes simplex virus (HSV)). SPCLs, as described, greatly aid all areas of drug discovery and research involving peptides. □

Received 31 July; accepted 16 September 1991.

1. Merrifield, R. B. *J. Am. chem. Soc.* **85**, 2149-2154 (1963).

2. Geysen, H. M., Meloen, R. H. & Barteling, S. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3998-4002 (1984).

- Houghten, R. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 5131-5135 (1985); US Patent 4,631,211.
- Fodor, S. P. A. *et al. Science* **251**, 767-773 (1991).
- Geysen, H. M., Rodda, S. J. & Mason, T. J. *Molec. Immun.* **23**, 709-715 (1986).
- Scott, J. K. & Smith, G. P. *Science* **249**, 386-390 (1990).
- Devlin, J. J., Pangniban, L. C. & Devlin, P. E. *Science* **249**, 404-406 (1990).
- Cwirla, S. E., Peters, E. A., Barrett, R. W. & Dower, W. J. *Proc. natn. Acad. Sci. U.S.A.* **87**, 6378-6382 (1990).
- Houghten, R. A., Hoffman, S. R. & Niman, H. L. *Vaccines 86* (eds Brown, F., Chanock, R. M. & Lerner, R. A.) 21-25 (Cold Spring Harbor Laboratory Press, New York, 1986).
- Appel, J. R., Pinilla, C., Niman, H. & Houghten, R. A. *J. Immun.* **144**, 976-983 (1990).
- Pinilla, C., Appel, J. R. & Houghten, R. A. *Peptides, Proc. 21st Eur. Peptide Symp.* (eds Giralt, E. & Andreu, D.) 860 (Escom Science, Leiden, the Netherlands, 1991).
- Houghten, R. A., Appel, J. R. & Pinilla, C. *Peptide Chemistry 1987* (eds Shiba, T. & Sakakibara, S.) 769-774 (Protein Research Foundation, Minoh-Shi, Osaka, 1987).
- Houghten, R. A., Appel, J. R. & Pinilla, C. *Vaccines 88* (eds Ginsberg, H., Brown, F., Lerner, R. & Chanock, R. M.) 9-12 (Cold Spring Harbor Laboratory Press, New York, 1988).
- Boman, H. G. & Hultmark, D. A. *Rev. Microbiol.* **41**, 103-126 (1987).
- Zaslloff, M. *Proc. natn. Acad. Sci. U.S.A.* **84**, 5449-5453 (1987).
- Seisted, M. E., Brown, D. M., DeLange, R. J. & Lehner, R. I. *J. Biol. Chem.* **258**, 14485-14489 (1983).
- Cuervo, J. H., Rodriguez, B. & Houghten, R. A. *Peptide Res.* **1**, 81-86 (1988).
- Blondelle, S. E. & Houghten, R. A. *Peptide Res.* **4**, 12-18 (1991).
- Blondelle, S. E. & Houghten, R. A. *Biochemistry* **30**, 4671-4678 (1991).
- Wilson, I. A. *et al. Cell* **37**, 767-778 (1984).
- Gisin, B. F. *Analytica chim. Acta* **58**, 248-249 (1972).
- Kaiser, E. T., Colescott, R. L., Bossinger, C. D. & Cook, P. I. *Analyt. Biochem.* **34**, 595-598 (1970).
- Tam, J. P., Heath, W. F. & Merrifield, R. B. *J. Am. chem. Soc.* **105**, 6442-6455 (1983).
- Houghten, R. A., Bray, M. K., DeGraw, S. T. & Kirby, C. J. *Int. J. Peptide Protein Res.* **27**, 673-678 (1986).

ACKNOWLEDGEMENTS. We thank P. Blanc, D. Burcin, S. McPherson, N. Salazar, L. Simpkins and J. Winkle for their technical assistance, E. Silva for manuscript preparation and H. Rapoport, S. Kent, and F. Bloom for their helpful comments and discussions. This work was funded by Irtex Pharmaceuticals, San Diego, California.

ERRATUM

Ancient oceans, ice sheets and the hydrological cycle on Mars

V. R. Baker^{*†}, R. G. Strom^{*}, V. C. Gulick^{*†‡},
J. S. Kargel^{*}, G. Komatsu^{*} & V. S. Kale[‡]

^{*} Lunar and Planetary Laboratory, University of Arizona, Tucson, Arizona 85721, USA

[†] Department of Geosciences, University of Arizona, Tucson, Arizona 85721, USA

[‡] Department of Geography, University of Poona, Pune 411 007, India

Nature **352**, 589-594 (1991).

IN this Article in the 15 August 1991 issue, an error in the *Nature* office led to the omission of a line from Table 3. In addition, some corrections noted by the authors were not made before publication. Corrected versions of the relevant passages appear below and in reprints.

TABLE 3 Possible sources of CO₂ related to catastrophic outflow and ocean formation on Mars

Source	CO ₂ partial pressure (mbar)
North polar cap	~20
Massive volcanism	~100
Adsorbed on regolith: ocean basin	≤100
land	≤350
Groundwater	≤1,300
CO ₂ clathrate	≤4,000

Fluvial history

In first paragraph:

The very high infiltration capacities of common martian surface rocks (lava flows and impact-brecciated regolith) would allow subsurface aquifers to be replenished easily, so that head differentials could be sustained and drive prolonged groundwater flow. The resulting sapping³¹ would then produce the observed pattern of structurally controlled, low-density valley networks²⁴.

Although Noachian valleys are consistent with atmospheric

water cycling, including precipitation, direct evidence for the associated oceanic sink would have been removed by subsequent erosion and deposition.

In final paragraph:

The time required to produce the assemblage of postulated martian glacial features formed by the Austral ice sheet is estimated from terrestrial rates to be 10^5 to 2×10^7 years.

Formation and dissipation of oceans

The less-common post-Noachian martian valley networks and widespread glacial features were formed at approximately the same time as subsurface water was catastrophically released to form the outflow channels. An explanation of their origin is suggested by the observed association of the largest outflow channels with great fracture zones³⁷ extending from the Tharsis bulge (Fig. 1). The fractures, it is suggested, served as immense flow conduits, the Memnonia Fossae conducting the flow west from Tharsis to Mangala Vallis³⁸, and the Valles Marineris system and parallel fractures conveying the flow eastward to Kasei and Maja Valles as well as to the channels of the Chryse trough (the Shalbatana, Simud, Tiu and Area valleys).

In third paragraph:

Sudden catastrophic discharges of water, forming outflow channels and chaotic terrain, could occur where the permafrost cap thinned to some critical threshold thickness (on average the melting geotherm in Tharsis would rise from a depth of 3 km to 1.5 km) or where fractures associated with magmatic intrusions propagated through the ground-water zone to the surface.

REFERENCES

31. Baker, V. R., Kochel, R. C., Laity, J. E. & Howard, A. D. in *Ground-water Geomorphology* (eds Higgins, C. G. & Coates, D. R.) 235–266 (Spec. Pap. 252, Geological Society of America, Boulder, Colorado, 1990).
34. Grant, J. A. III thesis, Brown Univ. (1990).

GUIDE TO AUTHORS

PLEASE follow these guidelines so that your manuscript may be handled expeditiously.

Nature is an international journal covering all the sciences. Contributors should therefore bear in mind those readers who work in other fields and those for whom English is a second language, and write clearly and simply, avoiding unnecessary technical terminology. Space in the journal is limited, making competition for publication severe. Brevity is highly valued. One printed page of *Nature*, without interruptions of the text, contains about 1,300 words.

Manuscripts are selected for publication according to editorial assessment of their originality and reports from independent referees. They can be sent to London or Washington and should be addressed to the Editor. Manuscripts may be dealt with in either office, depending on the subject matter, and will where necessary be sent between offices by overnight courier. All manuscripts are acknowledged on receipt but fewer than half are sent for review. Those that are not reviewed are returned as rapidly as possible so that they may be submitted elsewhere without delay. Contributors may suggest reviewers; limited requests for the exclusion of specific reviewers are usually heeded. Manuscripts are usually sent to two or three reviewers, who are chosen for their expertise rather than their geographical location. Manuscripts accepted for publication are typeset from the London office.

Nature requests authors to deposit sequence and crystallographic data in the databases that exist for this purpose, and to mention availability of these data.

Once a manuscript is accepted for publication, contributors will receive galley proofs in about 4 weeks. *Nature*'s staff will edit manuscripts with a view to brevity and clarity, so contributors should check their proofs carefully. Manuscripts are generally published 2–3 weeks after receipt of corrected proofs. *Nature* does not exact page charges. Contributors receive a reprint order form with their proofs; reprint orders are processed after the manuscript is published and payment received.

Categories of paper

Review articles survey recent developments in a field. Most are commissioned, but suggestions are welcome in the form of a one-page synopsis addressed to the Reviews Coordinator. Length is negotiable in advance.

Articles are research reports whose conclusions are of general interest and which are sufficiently rounded to be a substantial aid to understanding. They should not have more than 3,000 words of text (not including figure legends) or more than six display items (figures and tables) and should not occupy more than five pages of *Nature*.

Articles start with a heading of 50–80 words written to advertise their content in general terms, to which editors will pay particular attention. The heading does not usually contain numbers, abbreviations or measurements. The introduction to the study is contained in the first two or three paragraphs of the article, which also briefly summarizes its results and implications. Articles have fewer than 50 references and must contain a few subheadings of two or three words.

Letters are short reports of outstanding novel findings whose implications are general and important enough to be of interest to those outside the field. Letters should have 1,000 or fewer words of text and four or fewer display items. The first paragraph describes, in not more than 150 words and without the use of abbreviations, the background, rationale and chief conclusions of the study for the particular benefit of non-specialist readers. Letters do not have subheadings and contain fewer than 30 references.

Commentary articles deal with issues in, or arising from, research that are also of interest to readers outside research. Some are commissioned but suggestions can be made to the commentary editor in the form of a one-page synopsis. Commentaries are normally between one and four pages of *Nature*.

News and Views articles inform non-specialist readers about new scientific advances, sometimes in the form of a conference report. Most are commissioned but proposals can be made in advance to the

News and Views editor

Scientific Correspondence is for discussion of topical scientific matters, including those published in *Nature*, and for miscellaneous contributions. Priority is given to letters of fewer than 500 words and 5 references.

Preparation of manuscripts

All manuscripts should be typed, double-spaced, on one side of the paper only. An original and three copies are required, each accompanied by artwork. If photographs are included, four sets of originals are required; for line drawings, one set of originals and three good-quality photocopies are acceptable. Reference lists, figure legends and tables should all be on separate sheets, all of which should be double-spaced and numbered. Relevant manuscripts in press, or submitted for publication elsewhere, should be included with each copy of a submitted manuscript, and clearly marked as such. Revised and resubmitted manuscripts should also be clearly marked as such and labelled with their manuscript numbers.

Titles say what the paper is about with the minimum of technical terminology, and in fewer than 60 characters in the case of Articles and Letters. Active verbs, numerical values, abbreviations and punctuation are to be avoided. Titles should contain one or two key words for indexing purposes.

Abstracts should be marked individually and clearly with the author's name and, when known, the manuscript number. Ideally, no figure should be larger than 28 by 22 cm. Figures with several parts are to be avoided and are permitted only if the parts are closely related, either experimentally or logically. Unreferenced originals of photographs should be provided. Original artwork is returned when a manuscript cannot be published.

Protein/nucleotide sequences should ideally be in the three-letter and not the single-letter code for amino acids. One column width of *Nature* can accommodate 20 amino acids or 60 base pairs. Numbering of sequences should be in the left-hand margin only, with a single space between rows.

Suggestions for cover illustrations, with captions and labelled with manuscript number, are welcome.

Colour artwork. A charge of £500 per page is made as a contribution towards the cost of reproducing colour figures. Inability to pay these costs will not prevent the publication of essential colour figures if the circumstances are explained. Proofs of colour artwork may be sent to contributors under separate cover from their galley proofs.

Figure legends should not exceed 300 words and ideally should be shorter. The figure is described first, then, briefly, the method. Reference to a method published elsewhere is preferable to a full description. Methods are not described in the text.

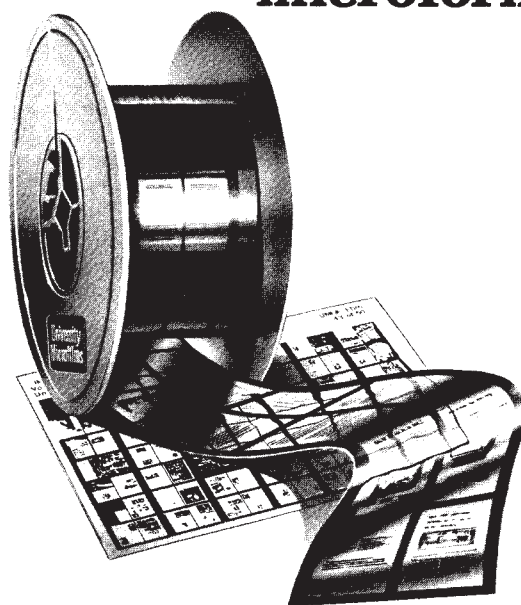
References are numbered sequentially as they appear in the text, followed by those in tables and finally by those in figure legends. Only papers published or in the press are numbered and included in the reference list. All other forms of reference, including unreferenced abstracts, should be cited in the text as a personal communication, manuscript submitted or in preparation. Text is not included in reference lists. References are abbreviated according to the *World List of Scientific Periodicals* (Butterworths, London, 1963–65). The first and last page numbers are included; reference to books should include publisher, place and date.

Abbreviations, symbols, units and Greek letters should be identified the first time they are used. Acronyms should be avoided whenever possible and, if used, defined. Footnotes are not used except for changed addresses.

Acknowledgements are brief; grant and contribution numbers are not allowed.

Submission. Manuscripts can be sent to the Editor at 4 Little Essex Street, London WC2R 3LF, UK or at 1137 National Press Building, Washington, DC 20045, USA. Manuscripts or proofs sent by air courier to London should be declared as 'manuscripts' and value \$5 to prevent the imposition of import duty and value-added tax (VAT).

nature is available in microform.



University Microfilms

International reproduces this publication in microform: microfiche and 16mm or 35mm film. For information about this publication or any of the more than 13,000 titles we offer, complete and mail the coupon to: University Microfilms International, 300 N. Zeeb Road, Ann Arbor, MI 48106. Call us toll-free for an immediate response: 800-521-3044. Or call collect in Michigan, Alaska and Hawaii: 313-761-4700.

☐ Please send information about these titles:

Name _____

Company/Institution _____

Address _____

City _____

State _____ Zip _____

Phone () _____

University
Microfilms
International

Databases for neuroscience

Steven L. Wertheim and Richard L. Sidman

Databases for neuroscience must handle large volumes of image and graphical data as well as domain knowledge. Several software development efforts are addressing these issues in an attempt to aid the acquisition, analysis and exchange of complex data sets.

WHILE neurophysiologists and behavioural neurobiologists have long used computers as essential tools for data collection and analysis, neuroanatomists have come to this technology more slowly as they work with high-resolution, memory-intensive image data. Only in the last few years has hardware capable of handling such data been available at reasonable cost. Two recent volumes review neuroanatomical data acquisition and analysis systems^{1,2}. The convenience, speed and precision of such systems have enhanced the number and complexity of neuroanatomical analyses that can be performed. The data produced are also relevant to scientists at the intersection of neuroanatomy with molecular neurobiology (who may be interested in the anatomical or developmental expression of particular gene products) and to clinicians (interested in localization of function and dysfunction in individual patients). The volume of neuroanatomical data being collected in electronic form, and the need to compare image data between experiments (or patients) and between researchers with overlapping interests, are leading neuroscientists to ask whether these data could be organized in widely accessible databases.

In parallel, government agencies have shown interest in funding a national project to create shared databases of brain structure, chemistry and function. The Institute of Medicine of the National Academy of Sciences, at the request of the National Institute of Mental Health (NIMH), has just completed a two-year study of the feasibility and desirability of such a project³.

Technology relevant to the construction of neuroscience databases is being developed in several laboratories, including ours. A guiding assumption is that an anatomical database is the scaffolding upon which to hang results from physiology, chemistry, pharmacology and behaviour. The technology has five aspects: (1) data acquisition and analysis (2) comparison to standards (3) data visualization and presentation (4) data retrieval and (5) communication.

Data acquisition and analysis are best understood, and most of the existing systems are devoted to these functions. They include image digitization, image processing, structure identification, cell counts, measurements of optical density

and activity, and measurements of area and volume. Systems vary according to whether they retain image data or only graphical abstractions (for example, contours and points) and the degree to which image processing is used to help identify relevant features^{4,5}.

Digital brain atlases

The ability to compare neuroanatomical results to a standard is a fundamental requirement for a neuroscience database. Differences between experimental animals, even within a species, differences in plane of section, histological variability, uncertainties regarding structure boundaries and inconsistent nomenclature all make simple verbal assertions (for example, 'Nucleus A projects to nucleus B' or 'Nucleus A contains substance x-like immunoreactivity'), and even drawings or selected photographs, insufficient for precise comparisons with other experimental results. In neuroanatomy, "standards" are represented by one or more brain atlases that have been created for each experimental species. Traditionally, anatomists present their data as handmade drawings ("charts") of two-dimensional (2D) sections through the brain with the positions of labelled cells, fibres, binding or activity indicated. Either the chart is an exact representation of the actual tissue section (thus being precise but non-standard and difficult to compare) or the results are transferred, using the researcher's best judgement of three-dimensional (3D) correspondence, to the plates of a well-known atlas (thus being standard, but imprecise).

Computer techniques can make such data transformations more precise and objective. They involve (1) the identification of corresponding landmarks in digital images of experimental and reference cross-sections and (2) a mathematical "warping" that transforms the experimental section into the reference section with the least possible distortion⁶. A system that combines a volumetric 3D atlas of the rat brain with data transformation techniques has been built by Rodgers and colleagues⁷. This system is also instructive in that it represents a seemingly simple, yet radical change in the way we use atlases. Rather than an abstract set of boundaries, their 3D atlas is a natural image data set that allows alignment and registration of exper-

imental image data. It thus takes a non-committal approach to the issue of structure boundaries, which may be the best way of reconciling the many maps that current staining technologies reveal.

Visualization tasks include 3D reconstruction, rendering (creating an image of the 3D model with assigned surface features, lighting and shading) and animation. Several university development projects in neuroscience have expended great effort on these techniques⁸⁻¹⁰. Today, advanced versions of this technology are available commercially either as complete applications (for example, Alias, Alias Research, Toronto) or as subroutine libraries (for example, Dore, Kubota Pacific Computer, Santa Clara, California).

Database issues

Data storage and retrieval is the domain of the database management system (DBMS). Although any organized information can broadly be called a "database", and even though the term is often used for any information in electronic form, only commercial DBMSs provide a sufficiently robust environment for the construction of a neuroscience database. The most modern DBMSs in widespread use are the relational systems, although object-oriented technology, the leading edge of development, is beginning to appear in commercial products.

For the past three years, we have been developing a laboratory and teaching tool for neuroscience with database and communication features^{11,12}. The current system, called NeuroDatabase, is now entering "field" testing. Our data are organized in the form of abstract tables managed from the earliest steps of data acquisition by a commercial relational DBMS (Oracle Corporation, Redwood City, California). The advantages of this approach are (1) the data structures are relatively easy to modify (2) the data themselves are resistant to corruption (3) the system will scale to handle very large data sets and (4) the software can be available for multi-user access and can be distributed across a network.

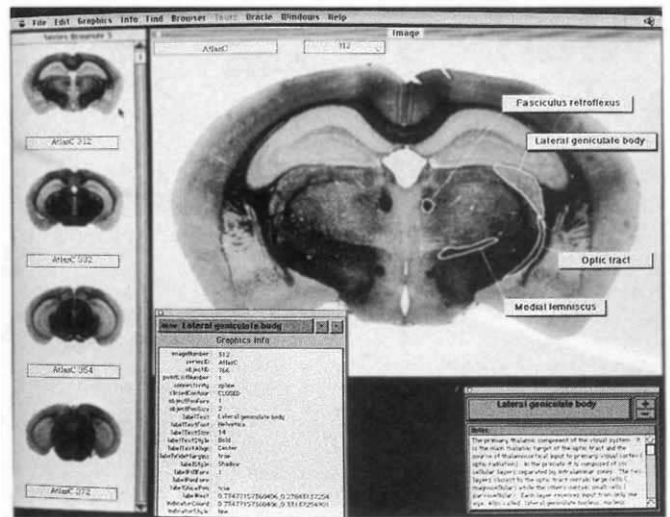
The demands of the subject matter have given NeuroDatabase two principal characteristics: a heavy emphasis on imaging and a provisional knowledge representation for neuroscience. Neuro-

Database allows one to manage many sets of images and their relationships. Examples of this might be two image series of the same tissue at different resolutions, or a set of image mosaics acquired with a computer-controlled microscope stage, video camera and frame-grabber board. We capture high-resolution images over large areas by repetitively moving the stage a precise amount and digitizing a frame to generate a mosaic built out of many individual tiles. Within limits, this technique can overcome the most severe disadvantages of both the video camera (low resolution) and the optical microscope (small field of view at high magnification). NeuroDatabase is designed to serve as the anatomical resource of first resort, in place of the actual microscope slides. It also allows the creation and flexible retrieval of 2D graphics associated with images such as boundary contours and positions of labelled cells (see figure). The DBMS allows linkages to be established between data in disparate data sets so that a user could, for example, call up all available *horizontal* images through the *lateral geniculate nucleus* in a *myelin stain* from *unoperated, adult cats*.

A rudimentary knowledge representation is necessary because the user and the software need not only to navigate through the elaborate terminology of the nervous system, but also through the relationships between terms. We use the term "knowledge representation" to cover vocabulary, data structures (for example, to contain the data from a particular type of experiment), object typing (for example, the putamen is a nucleus, dopamine is a neurotransmitter), hierarchical relationships between objects (for example, thalamus/lateral geniculate nucleus/lamina 3) and relationships between types (for example, a cell consists of cell body, axon, dendrites). Our knowledge representation is the minimum needed to aid creation and parsing of queries. A more elaborate knowledge representation is being constructed by Brinkley and colleagues¹³.

Another software package for the organization of research data is Brain Browser¹⁴. Brain Browser is useful for making a personal summary of results from the literature. Its card-file design can help one keep both verbal and graphical summaries of the anatomical information from large numbers of original articles. One can then quickly search the cards for all information relevant to a particular site or pathway. However, as it was created entirely within the HyperCard environment (Apple Corporation, Cupertino, California) and does not use a DBMS, it lacks the advantages mentioned above. And, as it is not linked to laboratory data acquisition,

A screen from the NeuroDatabase program. In the left window, image icons from a large series of mouse brain sections¹⁵ are accessible four at a time by use of the scroll bar. The first image is displayed at larger size in the central window. Functions available from the menu bar allow the user to request or create graphics on top of the image. Part of the database record for a graphic overlay is displayed in the lower left window. Glossary data on a structure are in the lower right window.



tion, it does not handle raw research data. These limitations are being remedied in a system under development by the authors of Brain Browser.

Communication and collaboration

Electronic communication of raw and analysed research data will be essential to the construction of neuroscience databases. The fact that each of the available neuroscience software packages, including ours, has its own data format, limits communication to users possessing identical systems. However, from an abstract point of view, the data structures within NeuroDatabase can be viewed as a specification for low-level data exchange that is independent of a particular software or hardware platform. In practice, NeuroDatabase provides mechanisms for the exchange of high-level data objects such as an image series with its associated graphics, or an experimental record. Thus, we can summarize the distributed nature of NeuroDatabase by saying that each laboratory has a private copy of the database to which its own data can be added. The vocabulary, data storage structures and analysis tools are shared. These common elements allow for electronic communication and collaboration.

To our knowledge, no single system, either commercially available or in development in universities, possesses all the features necessary to serve as a model for a large shared neuroscience resource. However, most of the necessary technology is available and at a recent workshop held by the NIMH, those participants who are actively developing software expressed a willingness to share technology so that their achievements need not be re-invented.

We are making NeuroDatabase available under a low-cost academic licence to interested laboratories who have the appropriate hardware and software.

Data collection and 2D analysis is done on a Macintosh II with 8-Mbyte RAM, hard disk and 19-inch, 8-bit colour display. Rewritable optical disks (Sony Corporation, San Jose, California) are used for archival data storage and data exchange. Three-dimensional reconstruction (optional) will be done on Silicon Graphics' Iris (Mountain View, California). A licence for the Oracle database is also required. □

Steven L. Wertheim and Richard L. Sidman are at the New England Regional Primate Research Center, Harvard Medical School, One Pine Hill Drive, Southborough, Massachusetts 01772-9102, USA. This work was partially supported by the NIH. For more information, fill in reader service number 100.

1. Capowski, J. J. *Computer Techniques in Neuroanatomy* (Plenum, New York, 1989).
2. *Three-dimensional Neuroimaging* (ed. Toga, A. W.) (Raven, New York, 1990).
3. *Mapping the Brain and its Functions* (eds Pechura, C. M. & Martin, J. B.) (National Academy, Washington, D.C., 1991).
4. Allen, B. A. & Levinthal, C. in *Three-dimensional Neuroimaging* (ed. Toga, A. W.) 211-233 (Raven, New York, 1990).
5. Brinkley, J. F. *Proc. 15th Annual Symp. on Computer Applications in Medical Care* (IEEE Computer Society, Los Alamitos, 1991) (in the press).
6. Bookstein, Fred L. in *Image Processing in Medical Imaging Lecture Notes in Computer Science*, Vol. 511 (eds Colchester, A. & Hawkes, D.) 326-342 (Springer, Berlin, 1991).
7. Rodgers, W. T., Moser, A. R. & Schwaber, J. S. in *Three-dimensional Neuroimaging* (ed. Toga, A. W.) 123-148 (Raven, New York, 1990).
8. Pearlstein, R. A. & Sidman, R. L. *Anal. Quant. Cytol. Histol.* 8(2), 89-95 (1986).
9. Prothero, J. S. & Prothero, J. W. *Proc. natn. Computer Graphics Assoc.* 1, 187-192 (1989).
10. Toga, A. W. & Payne, B. A. in *Three-dimensional Neuroimaging* (ed. Toga, A. W.) 251-271 (Raven, New York, 1990).
11. Wertheim, S. L. *Proc. 13th Annual Symp. on Computer Applications in Medical Care*, 399-404 (IEEE Computer Society, Los Alamitos, 1989).
12. Wertheim, S. L. *Proc. 14th Annual Symp. on Computer Applications in Medical Care*, 389-393 (IEEE Computer Society, Los Alamitos, 1990).
13. Eno, K., Sundsten, J. W. & Brinkley, J. F. *Proc. 15th Annual Symp. on Computer Applications in Medical Care* (in the press).
14. Bloom, F. E., Young, W. G. & Kim, Y. *The Brain Browser* (Academic, San Diego, 1989).
15. Sidman, R. L., Angevine, J. B. Jr & Taber-Pierce, E. *Atlas of the Mouse Brain and Spinal Cord* (Harvard University Press, Cambridge, 1971).

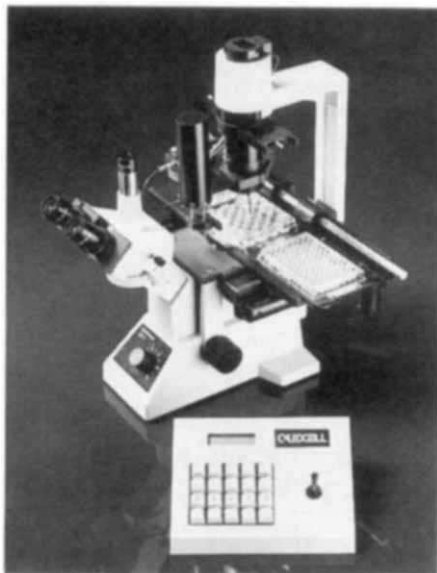
Neuroscience and all that jazz

New Orleans, Louisiana, is the venue for next week's 21st Annual Meeting of the Society for Neuroscience. New developments to watch out for include a single-cell transfer system and a selection of implantable transmitters for studying animals under field conditions.

A varied selection of **tracers for neuroscience research** will be featured in Molecular Probes' booth, number 952 in New Orleans (*Reader Service No. 101*). New products will include unsaturated versions of DiI and DiO, which diffuse faster within the membrane than earlier versions of these probes, and an assortment of fluorescent dextrans and fluorescent latex products, some of which are useful for retrograde tracing. Dialkylcarbocyanine fluorescent probes, particularly the octadecyl (C_{18}) indo- and oxacarbocyanines, which are commonly referred to as DiI and DiO, have become standard tools for neuronal tracing in both culture and fixed tissues. Their suitability for this application is based upon a combination of favorable optical and physical properties, the company says. DiI and DiO can be used both as retrograde and anterograde tracers as they diffuse laterally within the membrane. The diffusion process can, however, be accelerated by introducing unsaturation in the probes' alkyl tails. Two new products, Fast DiI and Fast DiO (di-unsaturated linoleyl DiI and DiO), have been shown by Dr Steve Senft of Washington University to undergo diffusion that is significantly faster than DiI or DiO. A new product in the fluorescent dextran line, "Fluoro Ruby", can be used for both anterograde and retrograde transport. To overcome the difficulty of detecting blue fluorescent latex in neurons, Molecular Probes has reconstituted its fluorescent latex kit for retrograde transport. The new kit contains a red fluorescent latex product in place of the blue fluorescent latex.

Cell manipulation

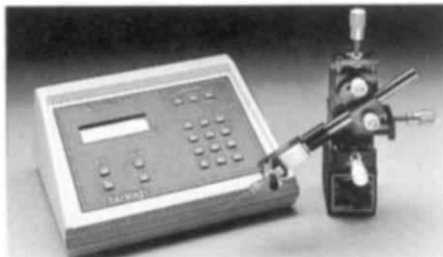
The Quixell **cell transfer system** from Stoelting uses automated micromanipulation to collect single cells, one cell at a time, in fine pipettes with exact precision and at a speed rivaling that of serial dilution techniques (*Reader Service No. 102*). A single cell is targeted, then, with a single keystroke, the cell is collected into the pipette and automatically transferred into the well of a microtitre tray. Each of a range of 128 "clicks", or button presses, moves a minute amount of fluid into or out of the tip of a glass micropipette. About ten clicks will draw in a cell that is located right at the pipette tip. If two cells are drawn in, one can be expelled with a couple of clicks on the "out" button, Stoelting says. The Quixell



Singling out cells is a cinch with Quixell, says Stoelting.

cell transfer system should be of interest to researchers isolating pure cell lines, generating cell cultures, using the polymerase chain reaction technology, producing monoclonal antibodies, or any task where cell selection is useful. It consists of a motorized microscope stage (with micrometre precision) that is capable of holding up to two microtitre trays, petri dishes or slides, and a robot arm that both holds and controls the pipetting system. See booths 127/129.

Burleigh Instruments has introduced a new **cell penetrator system** for generating ultra-fast micropipette speeds for applications such as intracellular recording and microinjection (*Reader Service No. 103*). The system uses a precision flexure assembly to produce straight, vibration-free motions of up to 70 μm , the company says. The flexure assembly is driven by a low-voltage piezoelectric element. The PZ-100 StepDriver controls pipette speed and position. It has programmable for-



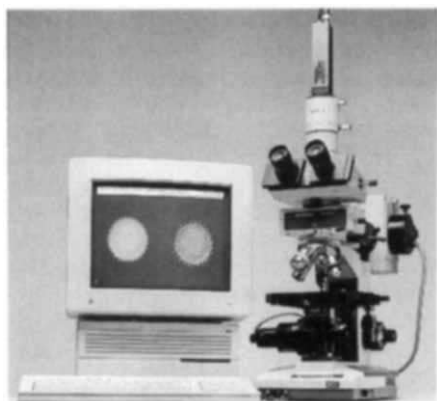
Burleigh's cell penetrator system.

ward and reverse step functions, an adjustable slew rate, smooth linear control and NVRAM set-up memory. Settings are in constant view on a backlit liquid-crystal display. An RS-232 communication port is fitted as standard. Burleigh Instruments says that the cell penetrator system can generate pipette speeds in excess of 300 $\mu\text{m ms}^{-1}$. Burleigh Instruments will be in booth 820.

Gadgets and gear

Mini-Mitter has introduced a new line of **implantable transmitters** for monitoring the temperature and heart rate of wild and domestic animals under field conditions (*Reader Service No. 104*). With a typical range of 1 km, the battery life of the transmitters is tailored to research design. The temperature transmitters have an accuracy of $\pm 0.2^\circ\text{C}$ over the life-time of the battery, Mini-Mitter says. The Datacol automated data acquisition system for use with these units collects data in real time from up to 50 unrestrained animals. The devices can also be used with multiply-housed animals. Stated applications include ovulation detection, energetics studies, disease assessment, drug responses, behavioural interaction studies, and the study of physiological stress resulting from confinement, relocation, noise, social interaction, heat/cold and other variables. For a demonstration—see booths 249/251.

Stop by Vaytek's booth, number 650, and see the new non-DSP board version of Micro-Tome — the **image processing system** that provides an inexpensive way to transform a standard wide-field microscope, video camera and an IBM AT/compatible computer into a digital confocal microscope (*Reader Service No. 105*). According to Vaytek, Micro-Tome offers several advantages over conventional laser scanning confocal microscopes: (1) any wavelength of light can be used (2) it works with transmitted light, DIC, polarized images or laser illumination (3) it will not burn the specimen (4) images can be captured at video rates (5) photobleaching is minimized and (6) it is easy to install. The non-DSP board version significantly reduces the price for Micro-Tome. Prices now start at \$995 (US). To produce confocal images, Micro-Tome uses published deconvolution algorithms to remove the out-of-focus haze from a microscope image. This set-up, Vaytek says, can be used to

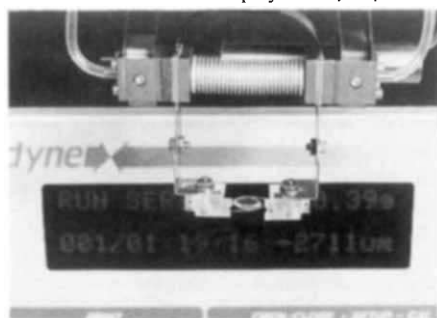


Vaytek's low-cost alternative to the laser scanning confocal microscope.

produce images comparable to those made by pinhole-based confocal microscopes. The non-DSP board version is part of an upgrade package for Micro-Tome to version 2.2. Additional features include: the ability to store and read point-spread functions; top-view projections; the option to use either theoretical or experimental point-spread functions; brightness, contrast and thresholding controls; virtual memory management; and variable slice space incrementing.

Brandel **receptor-binding harvesters**, available exclusively from Semat, are designed to reduce significantly the time and effort involved in sample preparation associated with thymidine incorporation studies, receptor-binding assays, receptor screening and other procedures requiring the filtering of samples from wells or test tubes (*Reader Service No. 106*). The harvesters can be used to process multiple samples simultaneously, separating cells or membranes from liquid and depositing them onto filter paper. During the procedure, wells or test tubes are washed thoroughly to flush out any remaining sample. Filtrate can be collected into a common reservoir, individual vials or test tubes. The harvesters are supplied in a variety of sizes and configurations for processing 12 to 96 samples. All models accept interchangeable harvesting probes. See Semat in booths 1312/1314/1316.

The Vitrodyne V-200 is a new **tabletop materials tester** from Liveco that can be used to evaluate the physical properties



Liveco's materials stress tester.

of a wide variety of biomaterials such as explanted tissues, biopolymers, implants and sensors (*Reader Service No. 107*). The device can be used to make ultrasensitive measurements within a controlled environment (that is, controlled for temperature, pH, salinity and humidity) and in a range not previously accessible with conventional testing apparatus, Liveco says. Commonly run tests include stress/strain, long-term creep, stress relaxation, flexural three-point bend, uniaxial tension and compression, and continuous static or dynamic programmes. The latest version includes a complete materials testing software package and a laser scan micrometer allowing displacement measurements in two axes simultaneously.

Alabama Research and Development is offering the Krumdieck **tissue slicer**, an automated microtome that has applications in a variety of scientific fields — biochemistry, physiology, pharmacology and toxicology (*Reader Service No. 108*). As it can be sterilized, the tissue slicer is suitable for the rapid preparation of aseptic slices of live tissues for organ culture. The tissue slicer can produce slices from approximately 60- μ m thickness and up to a maximum rate of one slice every three to four seconds, the company says. The tissue slicer has been designed to eliminate many of the major sources of error in tissue slice work, namely the use of slices of different and uneven thickness with irregular, non-reproducible damage at the cut surfaces.

Controlled injection

The NeuroPhore BH-2 system from Medical Systems is designed to facilitate the **controlled injection of fluids from micropipettes** (*Reader Service No. 109*). It features both pneumatic and iontophoretic pump modules. Using NeuroPhore BH-2, intracellular or extracellular ejections of minute volumes can be delivered using up to five pumps in parallel. In addition, the ejection routine for each pump can be independently programmed. The company also offers multi-barrel micropipettes for use with the NeuroPhore BH-2 system. This enables researchers to use the apparatus to eject several different fluid combinations to investigate cell responses. See booths 119/123.

The new List-Bogomoletz "Jumping Table" (L/M-CC-1000), which is available from List-electronic, allows the user to produce reproducible and **automatically-controlled applications of solutions to any isolated voltage-clamped cell or membrane patch** (*Reader Service No. 110*). The protocol includes step-like applications of a desired sequence of solutions (up to 29 in a series) and

ADVERTISEMENTS

Teflon[®] Coated and Bare Wire from MEDWIRE



Medwire provides the highest quality platinum, platinum alloy, silver and stainless steel wire in both single and multistrand configurations.

For prompt delivery or additional information, please contact:

MEDWIRE, an affiliate of Sigmund Cohn Corp.
121 So. Columbus Ave. Mount Vernon, NY 10553
Tel: (914) 664-5300 Fax: (914) 664-5377

*Reg. TM of E.I. DuPont & Co.

Reader Service No.25

Professional Custom Services

DNA SYNTHESIS PROTEIN SEQUENCING PEPTIDE SYNTHESIS

Rapid Service • Guaranteed Quality • Ample Quantity
Competitive Prices • Free Quotations
World-wide Delivery

V.G. Services 1200 Wonderland Rd. S. Bldg. #7, Unit #2
London, Ontario Canada N6L 1A8
CALL/FAX (519) 652-3758 ALT. FAX (519) 652-9622

Reader Service No.19

Custom Peptides

\$30 a residue, 10 μ mole scale
\$80 a residue, 100 μ mole scale

Research Genetics

1-800-533-4363

In the United Kingdom 0 800 69 1393

In Canada 1-800 533 4363

FAX 205-536-9016

Reader Service No.32



COMPLETE SYSTEMS FOR XENOPUS OOCYTE INJECTION

INT. SALES: +44. 984. 40226.

FAX: + 44. 984. 41166

Singer Instrument Co. Ltd.

Roadwater, Watchet, Somerset,
England, TA23 0QL. Tel: Washford (0984) 40226
Telex: 337492 Comcab G

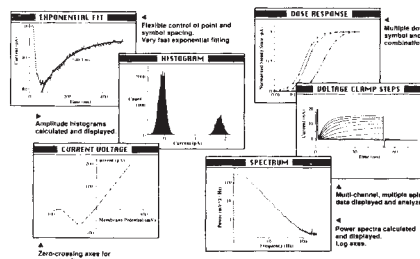
Reader Service No.50

synchronous with each application is measurement of membrane ionic current (or single-channel current) at the voltage clamp conditions. As the set-up is under computer control, only 5 to 20 minutes may be needed to obtain a complete "dose-response" curve for a certain substance. List-electronic says the procedure ("concentration" plus "voltage clamp") is not only applicable to neurons but also to glial cells, fibroblasts, oocytes, heart muscle cells, smooth muscle cells, blood cells and protozoa. See booths 547/549/551.

Software products

Axon Instruments enters the world of the Macintosh with its three new **software programs for data acquisition and analysis** — AxoData and AxoGraph for signal acquisition and analysis and AxoVideo for image acquisition and processing (*Reader Service No. 111*). The \$1,000 (US) AxoData program is intended for use in current-clamp, voltage-clamp and a variety of other experiments requiring concurrent waveform generation and sweep acquisition. This laboratory control and data acquisition program allows the user to create sophisticated stimulus protocols graphically and acquire sweeps of multi-channel data. AxoGraph, a \$500

(US) integrated graphics and analysis program specifically optimized to handle large data sets, provides a fast route from raw data to the production of publication-quality figures of electrophysiological and other data. Completing the line up is AxoVideo, a \$3,000 (US) program that automates the capture and processing of



AxoGraph — Axon Instruments' data analysis and graphics program for the Mac.

time-lapse video images. AxoVideo enables the interactive manipulation of individual images to produce publication-quality figures. A built-in programming language allows both customization of AxoVideo and automation of repetitive image processing tasks. Technical back-up by telephone is included in the price. See booths 1145/1147/1149/1151.

With its advanced multi-tasking design, the 1401 plus **laboratory interface** from Cambridge Electronic Design is intended for use in applications where data acquisition, experimental control and real-time processing are required simultaneously (*Reader Service No. 112*). The 1401 plus, a higher performance alternative to the CED 1401 model, retains many of the



1401 plus laboratory interface and Spike2 software upgrade — see booths 346/348.

features of the standard unit, but has the added power provided by a built-in 32-bit processor. Taking advantage of the improved hardware, software engineers at Cambridge Electronic have developed new versions of many of the company's life science applications packages. Several enhancements to the Spike2 software suite of programs will be demonstrated in New Orleans. The introduction of the 1401 plus has enabled the company to develop online routines for capturing single spikes from multi-unit recordings and sorting them using template matching machines. Online and offline sorting routines will be fully integrated into the Spike2 suite, allowing the resultant unit data to be processed using standard tools. Researchers wanting a sneak preview of the pre-release version of Spike2 for the Macintosh II should visit booths 346/348.

Antibody assortment

Four new **antibodies for Alzheimer research** will feature on the Boehringer stand, number 124, at the Society for Neuroscience meeting (*Reader Service No. 113*). Anti-Alz 90, which belongs to the IgG subclass, stains Alzheimer precursor protein pre 4 from rat and human. Function testing of the antibody on western blots using brain homogenates ensures performance of the antibody, Boehringer says. Anti- β -Amyloid, a polyclonal antibody raised against synthetic β -amyloid, stains western blots and senile plaques in cryo- and formalin-fixed brain sections. Anti-Choline Acetyltransferase reacts with choline acetyltransferase from rat, pig and human. Cryosections, Vibratome sections and cell lysates can be used as sample material. Detection is also possible on western blots. The monoclonal antibody Anti-GAP-43/B-50 binds to growth-associated protein from rat, cat, hamster and human. Growth-associated protein is expressed at elevated levels in developing and regenerating neurons during axon growth. The antibody can be used in immunohistochemical and western blot procedures.

Affiniti Research Products has introduced a new **rabbit polyclonal antibody to neuropeptide tyrosine (NPY)** for use in radioimmunoassays (*Reader Service No. 114*). With a detection sensitivity down to 0.5 femtomole per assay tube, the NPY antibody is known to react with rat, human and porcine NPY, but shows no crossreactivity with rat or human peptide with tyrosine (PYY), rat or human pancreatic polypeptide, glucagon, or with the NPY[19-36] fragment, Affiniti says.

The Finnish company, Biohit Oy, is distributing a new **mouse monoclonal antibody to the neurotransmitter GABA** (gamma-aminobutyric acid), which was developed by Locugenex (*Reader Service No. 115*). According to Locugenex, the monoclonal antibody to GABA is highly specific and eliminates the polyclonal background associated with antibodies used previously. The monoclonal antibody can be used to provide information on the co-localization of GABA with respect to other neurotransmitter substances studied in cells. It can be used with all commonly used cell research techniques, including fluorescence microscopy, and with frozen sections, Vibratome sections and paraffin-embedded sections. □

These notes are compiled by Diane Gershon from information provided by the manufacturers. To obtain further details, use the reader service card bound inside the journal. Prices quoted are sometimes nominal, and apply only within the country indicated.

ADVERTISEMENTS

ARE YOU LOOKING FOR A SOURCE OF TRITIATED DIHYDRO-DIDS ?

HSC RESEARCH & DEVELOPMENT
TORONTO, CANADA

TEL 416-598-5982
FAX 416-598-7163

SHIPPED WORLDWIDE

Reader Service No.16

A new label....



....unique products

Whether it's 25µL of a monoclonal antibody,
50pmol of an oligonucleotide probe or
5mg of a neurochemical.

Now they've no place to hide!

AFFINITI Research Products Ltd.,
10 Cossall Industrial Estate, Ilkeston, DE7 5UG, UK

Tel: (+44/0)602 442232;

Fax: (+44/0)602 442313

Specialising in Neuroscience Research

Reader Service No.17